



# Etude des grands assemblages protéolytiques de la famille TET : processus d'oligomérisation et régulation fonctionnelle associée

Alexandre Appolaire

## ► To cite this version:

Alexandre Appolaire. Etude des grands assemblages protéolytiques de la famille TET : processus d'oligomérisation et régulation fonctionnelle associée. Biologie structurale [q-bio.BM]. Université de Grenoble, 2014. Français. NNT : 2014GRENV062 . tel-01314147

**HAL Id: tel-01314147**

**<https://theses.hal.science/tel-01314147>**

Submitted on 10 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Chimie Biologie**

Arrêté ministériel : 7 août 2006

Présentée par

**Alexandre APPOLAIRE**

Thèse dirigée par **Bruno FRANZETTI** et **Frank GABEL**

préparée au sein du **Groupe Extremophiles et Grands Assemblages Moléculaires**

dans l'**École Doctorale de Chimie et Science du Vivant**

## Étude des grands assemblages protéolytiques de la famille TET : processus d'oligomérisation et régulation fonctionnelle associée.

Thèse soutenue publiquement le 15 décembre 2014,  
devant le jury composé de :

**Prof. Bertrand FRIGUET**  
(Université Pierre et Marie Curie)

**Rapporteur**

**Dr. Vincent VILLERET**  
(Université Lille 1)

**Rapporteur**

**Dr. Véronique RECEVEUR-BRECHOT**  
(Université Aix-Marseille)

**Examinatrice**

**Dr. Jean-François HERNANDEZ**  
(Université Montpellier 1)

**Examineur**

**Dr. Didier FLAMENT**  
(Université de Bretagne Occidentale)

**Examineur**





# ETUDE DES RELATIONS ENTRE PROCESSUS D'OLIGOMERISATION ET FONCTION DES AMINOPEPTIDASES TET

## RESUME

---

La protéolyse est une fonction clé de la cellule pour le maintien du protéome et pour la régulation de nombreux processus physiologiques. Le travail présenté dans cette thèse porte sur une famille de peptidases cytosoliques auto-compartmentés et énergie indépendantes découvertes chez les Archées, les aminopeptidases TET. Chez l'Archée hyperthermophile *Pyrococcus horikoshii*, organisme modèle de cette étude, il existe 3 TET présentant chacune des spécificités de substrats différentes. Les caractérisations structurales des différents membres connus de cette famille de peptidase ont révélé un assemblage dodécamérique creux en forme de tétraèdre d'environ 450 kDa. Des études récentes ont montré l'existence de complexes adoptant la même conformation que les TET dans les 3 domaines du vivant. La première partie du travail présenté vise à identifier des marqueurs structuraux caractéristiques de l'assemblage tétraédrique afin de déterminer sans ambiguïté l'appartenance de ces complexes à la famille des TET. La seconde partie de l'étude vise à élucider la question de la multiplicité des TET chez les Archées hyperthermophile mise en évidence grâce à une étude phylogénétique initiée pendant la thèse. L'étude en co-expression de PhTET2 et PhTET3 révèle que ces aminopeptidases sont capable de former un hétéro-oligomère présentant une activité enzymatique accrue vis-à-vis des homo-oligomères. La dernière partie du travail porte sur les relations oligomérisation-fonction d'un tel édifice. L'étude d'un mutant de l'oligomérisation de PhTET2 via une stratégie intégrative alliant biochimie, enzymologie, biophysique et des études in vivo ont permis de mettre en évidence un processus d'assemblage contrôlé permettant d'augmenter l'efficacité de la peptidase sur des peptides plus long. Enfin la méthode de variation de contraste en diffusion de neutrons aux petits angles appliqué à l'étude de l'hétéro-oligomère a permis de révéler une



topologie rationalise du complexe hétéro-oligomérique favorisant la formations de poches multi-catalytique. L'ensemble de ce travail vise à mieux comprendre l'intérêt biologique de telles structures dans la cellule.

Proteolysis is a very key function in the cell for the maintenance of the proteome and for the regulation of many physiological processes. The work here presented treats about a family of self-compartmented energy-independent cytosolic peptidases discovered in Archaea, the TET aminopeptidases. Three different TET showing different enzymatic specificities are expressed in the hyperthermophilic archaeon *Pyrococcus horikoshii*, which is the model organism for this study. The structural characterization of the known members of this family shows that they assemble with a very original structure forming dodecameric hollow tetrahedral particles of 450 kDa. Recent studies have revealed the existence of complexes that adopt the same conformation in the three domains of life. The first part of this work is aim to identify some structural markers to assigned without any ambiguity these peptidases to the TET family. The second part of this study is aim to better understand the multiplicity of TET peptidases in hyperthermophilic archaeon which has been highlighted by a phylogenomic study started during the PhD. The co-expression of PhTET2 and PhTET3 revealed that they can form a hetero-oligomeric complex with an enhanced enzymatic activity compared to the homo-oligomers. The last part of this work treats about the oligomerization-function relationship in such a structure. A mutagenesis strategy was used to slow down the oligomerization process of PhTET2, and, using an integrative strategy combining biochemistry, enzymology, biophysics and in vivo studies we were able to show that the oligomerization of the particle is a non-random controlled process aim to enhance the activity of the peptidase towards longer peptides. At last, the contrast variation technique in small angle neutron scattering allowed us to unravel the rational topology of the hetero-oligomer which favored the formation of multi-catalytic enzymatic pockets in the complex. This work is aim to better understand the biological interest of such a particle in the cell.

### REMERCIEMENTS

---

D'abord, merci à toi lecteur de t'attarder sur ce manuscrit qui regroupe mes travaux de thèse réalisés entre 2010 et 2014 dans le laboratoire « Extremophiles and Large Molecular Assemblies » ou ELMA, à l'institut de biologie structurale à Grenoble.

Je voudrai remercier Bruno Franzetti et Frank Gabel, mes directeurs de thèse, pour m'avoir fait confiance, m'avoir guidé et soutenu tout au long de ce projet. Merci aussi à Éric Girard pour toute son aide durant cette thèse. Cela aura été un réel plaisir de travailler avec vous ! Merci aussi à Jo Zaccai qui m'a permis de m'initier au monde merveilleux des petits angles.

Merci à tout le laboratoire ELMA, passé et présent. En particulier Asùn, pour les longues soirées passé au laboratoire à tenter de ~~conquérir le monde~~ montrer une activité au dimère de TET ; à Eva pour tous ce qu'elle m'a appris, de l'AUC à l'utilisation des FPLC en passant par le SAXS ; à Arnaud qui m'a épaulé durant mes premiers pas en tant qu'apprenti chercheur. Merci à Pierre et Domi pour la balance toujours bien salée.

Merci aussi à tous les collaborateurs qui ont participé à ce projet : Michael Haertlein et Martine Moulin pour les cellules deutérées ; Simonetta Gribaldo pour la merveilleuse histoire des TET ; à Guy Schoehn pour la microscopie électronique ; à Didier Flament et Véronique Rossi pour leur aide précieuse ; à Mohamed Jebbar ; à Phil Callow et Anne Martel pour leur aide sur D22 ; à Anne Godfroy pour les cellules de *Pyrococcus*.

Je voudrais ensuite remercier tout particulièrement ma voisine de bureau pendant ces 5 années à l'IBS, l'irremplaçable Mylène Ferruit (ne pas prononcer Ferruite). Beaucoup de choses à dire et peu de place... Merci... Merci pour tout =) Merci pour les dolipranes (je t'en doit quelques boîtes d'ailleurs), les cafés du matin dans le bureau de Chady et d'Eugénie, ou ceux de l'après-midi bien qu'ils aient été moins revigorant ;p, merci pour ton aide précieuse tout au long de ce projet, pour les innombrables commandes passées, pour les parties de pétanques, pour les parties de base-ball labo aussi, pour tous les tests d'activité, pour l'hétéro-oligomère, les appels du suicide, pour les séances de light

## Remerciements

---

painting, les shoots photos, les randos champignon sans champignon, les merveilleuses chansons d'alpes 1, etc... Bref, restes celle que tu es !

Merci à mes autres compères de bureau à l'IBS 1 comme ceux de l'IBS 2. Louise, merci pour toute ton aide, pour toutes les discussions et interrogations scientifiques ou pas, pour avoir participé à ma renommée de pâtissier, merci d'avoir gardé Gabriel aussi =)  
Merci Matteo pour toute ton aide en cristallographie et pour toutes les discussions de bureau souvent fructueuses ! Merci à tous les thésards que j'ai pu rencontrer pendant ces années à l'IBS et avec qui j'ai pu échanger : Yann Fichou, FX, Chady, ALF, Mathieu, Wiwi, Sylvain, Hugo, Eric ...

Un grand merci à tous le personnel de l'IBS qui m'a aidé tout au long de ce séjour parmi eux : Sylvie, Dominique R, Alain G, Blandine, Alain R ...

De gros bisous à ceux avec qui j'ai pu déjeuner pendant ces quelques années : Céline, Delphine, Isabelle, Aurianne, Hugues, Sonja, Maria ... Et à tous ceux que j'ai pu croiser dans l'aile 6 ou au 3<sup>ème</sup> étage par la suite !

Merci aussi à tous ceux que j'ai pu rencontrer au gré de mes pérégrinations Grenobloise et qui ont rendu ce séjour à Grenoble simplement fantastique ! Bisous au grand-duc versaillais alias Yann Huon de Kermadec, éternel amateur de poker ! Bisous à Guillaume Monin ! Bisous à Gildas Mainsant ! Gros bisous à la coloc du 93 cours de la libération ! Alice, Axel, Jo ! Bisous à la coloc Casimir Brenier aussi ! Zouki ! Solenne ! Alex ! Gabi ! Les apéros chez Sarah me manqueront et les soirées chez Gabi aussi !!

Des énormes bisous à Romain Talon, je n'oublierai pas notre victoire au concours photo ! Au bout de chaque rue une montagne !! Merci pour les russes blancs et les soirées à Mison ! Gros bisous à Marine ! Enormes bisous à Guénolé aussi, même si tu auras mis du temps à te souvenir de mon nom, les cafés de l'après-midi n'auront pas le même goût sans toi ! Gros bisous à la belle Sarah Ancelet ! Merci pour toutes les soirées passées à danser, à boire aussi et pour tous les fous-rire que j'ai pu avoir en ta compagnie ! Gros bisous à Michel et Marlène et au petit poulet alias les M&M&M's ! Merci pour les parties d'Aladdin ou de Bomberman calé dans le canapé ! Merci pour les champignons aussi ! Maouw !!

## Remerciements

---

Je voudrai aussi faire un bisou à ma maman, qui m'a toujours soutenu dans toutes les épreuves ! Merci aussi à ma tante Hélène et mes oncles Alphonse et Charlot ainsi qu'à mes frères et sœur Stéphane, Nadège et Gogo ! Gros bisous aussi à Bernard, Fred et Frank !

Enfin, je voudrai remercier ma future femme, Claire, sans toi cette thèse n'aurait pas été la même, tu m'auras soutenu et encouragé même quand rien n'allait plus. Merci à toi d'être là.



Merci aussi au petit Gabriel d'avoir pointé le bout de son nez pendant la dernière année de cette thèse, tu n'auras rendu la fin de cette thèse que plus palpitante !

*À mon père,*



## TABLE DES MATIERES

<b>Etude des relations entre processus d'oligomérisation et fonction des aminopeptidases TET</b>	<b>II</b>
Résumé	II
Remerciements	IV
<b>Tables des illustrations</b>	<b>XII</b>
Figures	XII
Tableaux	XIII
<b>Abréviations</b>	<b>XIV</b>
<b>1. Introduction</b>	<b>1</b>
1.1 Importance des peptidases cytosoliques dans les processus cellulaires	1
1.2 Classification des peptidases	2
1.2.1 Nomenclature des sites actifs	5
1.3 Peptidases oligomériques auto-compartmentées ATP-indépendantes	5
1.3.1 Peptidase Tricorn	5
1.3.2 Tripeptidyl peptidase II	8
1.3.3 DegP / HtrA	11
1.3.4 Bloemycin hydrolase / Gal6	14
1.3.5 Trilobed peptidase / PhAAP	17
1.3.6 DppA	20
1.3.7 PhP1 / PfP1	21
1.3.8 Pab87	22
1.4 Synthèse	24
1.5 Les aminopeptidases TET de <i>Pyrococcus horikoshii</i>	27
1.5.1 Historique de la découverte des aminopeptidases TET.	27
1.5.2 Fonction	28
1.5.3 Mécanisme enzymatique	28
1.5.4 Substrats	30
1.5.5 Co-facteurs et paramètres enzymatiques optimaux	31

## Sommaire

---

1.5.6	Structures	32
1.5.7	Mode d'action	34
1.5.8	Rôle physiologique	35
<b>1.6</b>	<b>Les Archées: organismes modèles pour les études des peptidases TET</b>	<b>36</b>
1.6.1	Le troisième domaine du vivant	36
1.6.2	<i>Pyrococcus horikoshii</i> , organisme modèle de la thèse	43
<b>1.7</b>	<b>Les peptidases cytosoliques de <i>Pyrococcus horikoshii</i></b>	<b>47</b>
<b>1.8</b>	<b>Objectifs et méthodologie de la thèse</b>	<b>55</b>
<b>2.</b>	<b>Méthodes utilisées</b>	<b>57</b>
<b>2.1</b>	<b>Biochimie</b>	<b>57</b>
2.1.1	Purification des protéines	57
2.1.2	Enzymologie	61
<b>2.2</b>	<b>Analyse structurale</b>	<b>65</b>
<b>2.3</b>	<b>Biophysique</b>	<b>67</b>
2.3.1	Diffusion aux petits angles	67
2.3.2	Ultracentrifugation analytique	73
3.2.1	Cristallisation des protéines	76
<b>3.</b>	<b>Identification de motifs structuraux conservés pour l'assignation des peptidases TET dans les génomes Archée</b>	<b>79</b>
<b>3.1</b>	<b>Introduction</b>	<b>79</b>
<b>3.2</b>	<b>Le domaine d'oligomérisation des TET</b>	<b>81</b>
<b>3.3</b>	<b>Les interfaces d'oligomérisations</b>	<b>82</b>
3.3.1	Interface de dimérisation	83
3.3.2	Interface de dodécamérisation	86
3.3.3	Cas particulier de Tm1048	88
<b>3.4</b>	<b>Déterminants structuraux spécifiques de la fonction des peptidases TET</b>	<b>90</b>
3.4.1	Les glycines : pivots de la structure	90
3.4.2	Conservation du site actif	90
<b>3.5</b>	<b>Synthèse : les motifs structuraux pour la détermination des complexes TET</b>	<b>91</b>
<b>3.6</b>	<b>Phylogénie des TET chez les Archées</b>	<b>94</b>
<b>3.7</b>	<b>Discussion</b>	<b>97</b>

<b>4. Les Aminopeptidases TET2 et TET3 de <i>Pyrococcus horikoshii</i> forment un peptidasome à sous-unités multiples doté de propriétés de destruction des polypeptides améliorées.</b>	<b>99</b>
4.1 Présentation du travail	99
4.2 "The TET2 and TET3 aminopeptidases from <i>Pyrococcus horikoshii</i> form a hetero-subunit peptidasome with enhanced peptide destruction properties".	102
4.3 Discussion et perspectives	115
<b>5. Assemblage des aminopeptidases TET : processus et signification biologique.</b>	<b>117</b>
5.1 Présentation du travail:	117
5.2 <i>Pyrococcus horikoshii</i> TET2 peptidase assembling process and associated functional regulation	121
5.3 Discussion et perspectives:	137
<b>6. Etude de l'assemblage de l'hétéro-oligomère des aminopeptidases TET2 et TET3 de <i>Pyrococcus horikoshii</i> par diffusion de neutrons aux petits angles.</b>	<b>139</b>
6.1 Présentation du travail	139
6.2 Small angle neutron scattering reveals the assembling mode and oligomeric organization of TET, a large dodecameric aminopeptidase	142
6.3 Discussion et perspectives :	162
<b>7. Résultats préliminaires de la caractérisation de PhTET4</b>	<b>165</b>
7.1 Identification	165
7.2 Purification	166
7.3 Microscopie électronique	167
7.4 Test enzymatique	168
7.5 Cristallisation	169
7.6 Discussion et perspectives :	170
<b>8 Conclusions générales et perspectives</b>	<b>171</b>
<b>Bibliographie</b>	<b>174</b>





## TABLES DES ILLUSTRATIONS

### FIGURES

---

Figure 1-1 : Représentation schématique de la dégradation séquentielle des protéines chez <i>Thermoplasma acidophilum</i> . ....	6
Figure 1-2 : Structure de la peptidase Tricorn. ....	7
Figure 1-3 : Représentation de l'iso surface de la capsid de Tricorn à une résolution de 1.3 nm. ....	7
Figure 1-4 : Structure de la peptidase TPPII. ....	9
Figure 1-5 : Structure hybride de l'hélice de TPPII. ....	10
Figure 1-6 : Structure et assemblages des formes inactives de la peptidase DegP. ....	12
Figure 1-7 : Détails de la structure du complexe à 24 sous-unités de DegP. ....	14
Figure 1-8 : Structure cristallographique de la peptidase Gal6. ....	15
Figure 1-9 : Représentation schématique des sites actifs de Gal6. ....	16
Figure 1-10 : Structure du monomère de PhAAP. ....	18
Figure 1-11 : Structure de l'hexamère de PhAAP. ....	19
Figure 1-12 : Structure de DppA. ....	21
Figure 1-13 : Structure de la peptidase Protease I de <i>Pyrococcus horikoshii</i> . ....	22
Figure 1-14 : Structure de la peptidase Pab87. ....	23
Figure 1-15 : Structures cristallographiques des peptidases formant de grands assemblages oligomériques représentés en ruban. ....	26
Figure 1-16 : Mécanisme catalytique des aminopeptidases TET. ....	29
Figure 1-17 : Structure cristallographique de PhTET1. ....	33
Figure 1-18 : Intérieur de la particule PhTET1. ....	35
Figure 1-19 : L'arbre phylogénétique universel du vivant proposé par Woese et col. ....	37
Figure 1-20 : Arbre phylogénétique et classification des Archées proposée par Brochier-Armanet et col. (2011). ....	41
Figure 1-21 : Photo d'un écosystème hydrothermal abyssal (IFREMER/VICTOR 6000). ....	42
Figure 1-22 : Diversité des peptidases cytosoliques chez <i>Pyrococcus horikoshii</i> . ....	54
Figure 2-1 : Densité de longueur de diffusion des principaux types de macromolécules biologiques en fonction de la concentration de D <sub>2</sub> O du solvant. ....	72
Figure 2-2 : Illustration schématique de la variation de contraste appliquée à un complexe protéique dont les monomères ont la même conformation. ....	73
Figure 3-1 : Des TET existeraient dans les trois domaines du vivant. ....	80
Figure 3-2 : Structure cristallographique d'AAP. ....	81
Figure 3-3 : Interface de dimérisation des TET. ....	84

## Tables des illustrations

---

Figure 3-4 : Structure cristallographique de Q11Z05_CYTH3. ....	185
Figure 3-5 : Interface de dodécamérisation des TET. ....	87
Figure 3-6 : Comparaison des structures cristallographiques de PhTET2 et TmPep1050. ....	88
Figure 3-7 : Superposition des structures cristallographiques de PhTET2.....	89
Figure 3-8 : Superposition des monomères de PhTET et agrandissement des zones où se trouvent les glycines conservées. ....	90
Figure 3-9 : Superposition des structures conservées du site actif de plusieurs TET.....	91
Figure 3-10 : Interfaces d'oligomérisation des TET. ....	93
Figure 3-11 : Alignement des séquences d'aminopeptidases TET. ....	94
Figure 3-12 : Représentation schématique de la phylogénie des TET et TRI chez les Archées. ....	96
Figure 7-1 : Alignement multiple des séquences des TET de <i>Pyrococcus horikoshii</i> réalisé avec clustalW2. .	166
Figure 7-2 : Purification et caractérisation de l'aminopeptidase PhTET4. ....	168

## TABLEAUX

---

Tableau 1-1 : Classification hiérarchique de la peptidase TET2 de <i>Pyrococcus horikoshii</i> . Le clan et la famille auxquels appartiennent TET2 sont indiqués par une flèche. ....	4
Tableau 1-2 : Résumé de différentes caractéristiques des peptidases présentées dans cette partie. ....	26
Tableau 3-1 : Structures identifiées lors de la recherche d'homologues structuraux de l'aminopeptidase PhTET2 à l'aide de l'outil Dali. * Présence uniquement du dimère dans l'unité asymétrique. ....	83

### ABREVIATIONS

ATP	Adénosine tri-phosphate
AAA-ATPase	ATPases associées à diverses activités cellulaires
VIH	Virus de l'immunodéficience humaine
PDB	<i>Protein Data Bank</i>
pNA	Para nitroaniline
AMC	Aminol méthyl coumarine
N-ter	N-terminale
C-ter	C-terminale
CHES	N-Cyclohexyl-2-aminoethanesulfonic acid
PIPES	Piperazine-N,N'-bis(2-ethanesulfonic acid)
Tris	Tris(hydroxymethyl)aminomethane
TAPS	N-tris(Hydroxymethyl)methyl-3-aminopropanesulfonic acid
EDTA	Ethylène Diamine Tétra Acétique
ARN	Acide ribonucléique
ADN	Acide désoxyribonucléique
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
DEAE	Diéthylaminoéthyl
$\mu$ RPC	<i>Micro Reverse Phase Column</i>
HPLC	<i>High pressure liquid chromatography</i>



# 1. INTRODUCTION

## 1.1 IMPORTANCE DES PEPTIDASES CYTOSOLIQUES DANS LES PROCESSUS CELLULAIRES

---

La protéolyse est une fonction essentielle de la cellule. Elle détermine le temps de demi-vie des protéines et permet de débarrasser la cellule des protéines déficientes induites par un stress ou un défaut de traduction (Goldberg, 2003; Breusing *et col.*, 2008). La protéolyse est impliquée dans le métabolisme, participe à l'autophagie et représente une source d'énergie dans les situations de carence. La destruction ciblée de facteurs de régulation permet également le contrôle direct de nombreux processus cellulaires (Varshavsky, 2005). Enfin, la protéolyse permet la maturation post traductionnelle et l'activation de nombreuses protéines ou peptides possédant une activité physiologique. La protéolyse est une fonction très régulée, impliquée dans les processus de vieillissement, de pathogénicité et dont l'altération est associée à de nombreuses maladies (Ingmer *et col.*, 2009; Chondrogianni *et col.*, 2010; Craik *et col.*, 2011).

La destruction des protéines dans le cytoplasme met en jeu des complexes protéolytiques énergie-dépendant comme le protéasome (Baumeister *et col.*, 1998), Lon (Cha *et col.*), HslVU (Rohrwild *et col.*, 1996) ou encore ClpP (Wang *et col.*, 1997). Il s'agit de complexes porteurs d'activités endopeptidases. Afin de protéger le protéome de la cellule de ces activités, un contrôle de ces dernières dans le temps et l'espace est indispensable. Les sites actifs sont ainsi situés à l'intérieur d'édifices multimériques creux. Cette auto-compartmentation permet une régulation *via* des modules de dépliement portés par des complexes AAA-ATPase ou des domaines de la protéine (Sauer *et col.*, 2011). Le rôle et le mode d'action de ces machines de destruction sont très étudiés (Sauer *et col.*, 2011). Cependant le cytosol contient aussi un grand nombre d'enzymes protéolytiques ou peptidases énergie-indépendantes qui représentent à elles seules environ 5% de notre génome selon la base de données des peptidases MEROPS (Rawlings *et col.*, 2014). Ces peptidases jouent des rôles clés dans des fonctions très variées. Elles interviennent dans les dernières étapes de la destruction des protéines en dégradant les polypeptides issus de l'action des endopeptidases ATP dépendantes (Yao *et col.*, 1999). Elles contribuent également au métabolisme pour l'anabolisme ou la

production d'énergie à partir d'acides aminés. Longtemps considérées comme des enzymes "de ménage" on réalise actuellement que les peptidase cytosoliques jouent également des rôles physiologiques très précis. Par exemple, la peptidase ERAAP (Endoplasmic Reticulum Aminopeptidase Associated with Antigen Processing) est responsable de l'hydrolyse de la partie N-terminale de peptides afin de les « personnaliser » en vue de leur présentation par le complexe majeur d'histocompatibilité I (van Endert, 2011). La peptidase Rénine est impliquée dans la régulation du renin-angiotensin system (RAS) dans lequel elle participe à la maturation du précurseur angiotensinogène qui jouera un rôle dans la régulation de la pression sanguine (Nurun Nabi *et col.*, 2013). La peptidase HIV 2 rétropepsine est produite par le virus du VIH et permet la maturation de précurseurs importants pour la propagation du virus (Goldfarb *et col.*, 2013). Enfin la thimet oligopeptidase joue un rôle dans la transduction de signal (Cunha *et col.*, 2008). Ces quelques exemples montrent que les peptidases cytosoliques et le peptidome qui résultent de leurs activités représentent des acteurs importants de la physiologie cellulaire. Les peptidases sont ainsi devenues des cibles importantes pour combattre de nombreuses maladies. Cependant, chez les eucaryotes, beaucoup d'entre elles n'ont pas été caractérisées. Ceci est dû à la complexité et à la diversité de ce type d'enzyme. Chez les procaryotes, les fonctions spécifiques des peptidases cytosoliques ont été très peu étudiées. Parmi ces peptidases, certaines sont capables de s'oligomériser et de former des grands assemblages moléculaires auto-compartmentés. De par leur taille et leurs organisations spatiales, elles sont susceptibles de faire l'objet de régulations, et de ce fait, d'être des acteurs importants de la dynamique des peptides intracellulaires. L'étude de ces grands assemblages et de leur lien avec l'adaptation aux conditions extrêmes ou à des processus associés à des pathologies représente un des principaux sujets d'étude du laboratoire « Extremophiles and Large Molecular Assemblies » qui m'a accueilli durant ma thèse.

### 1.2 CLASSIFICATION DES PEPTIDASES

---

La nomenclature des protéines, et en particulier des peptidases, est souvent source de confusion, ceci est notamment dû à la liberté laissée aux chercheurs pour nommer l'objet de leur travail. Cette situation amène parfois à obtenir plusieurs dénominations différentes pour un seul et même objet.

Pour définir les enzymes protéolytiques, on parle de protéases, de protéinases ou encore de peptidases. Ces termes ont des significations différentes. Ainsi, une protéase est une enzyme capable de dégrader une protéine en hydrolysant les liaisons peptidiques liant les résidus (Barrett *et col.*, 1986). La définition d'une peptidase est la même que celle d'une protéase mais généralisée à tout type de peptides, c'est le terme dont l'utilisation est recommandée par le « Nomenclature Committee of the International Union of Biochemistry and Molecular Biology » (NC-IUBMB) pour désigner toute enzyme capable d'hydrolyser une liaison peptidique. Le terme peptidase, associé à des suffixes ou préfixes permet de définir plus précisément l'activité d'une enzyme et de faciliter la communication entre chercheurs. Par exemple, une exopeptidase clive un ou plusieurs résidus à l'extrémité N-terminale ou C-terminale d'une chaîne polypeptidique ; une endopeptidase est capable de cliver à l'intérieur de la chaîne polypeptidique. Une exopeptidase qui clive uniquement à l'extrémité N-terminale un seul résidu est une aminopeptidase, si elle clive un di-ou tripeptide, c'est une dipeptidyl-ou tripeptidyl-peptidase. Dans le cas de l'extrémité C-terminale, on parle respectivement de carboxypeptidase, de peptidyl-dipeptidase ou de peptidyl-tripeptidase (handbook of proteolytic enzymes).

La base de données MEROPS répertorie toutes les peptidases ayant été caractérisées ainsi que les séquences de peptidases hypothétiques (<http://merops.sanger.ac.uk>). Les peptidases répertoriées dans MEROPS ainsi que dans le « handbook of proteolytic enzymes » sont classées selon un système hiérarchique, d'abord par rapport au type d'activité catalytique, puis par clan, par famille et enfin par peptidase.

Plusieurs types catalytiques existent, ils sont définis par la nature chimique des groupes responsables de la catalyse. À chaque type catalytique est associée une lettre afin de faciliter la nomenclature : peptidase Aspartique (A), à Cystéine (C), Glutamique (G), à Sérine (S), à Thréonine (T), les Metallopeptidases (M) et les peptidases non classés (U). Un septième type catalytique a récemment été découvert, il s'agit des lyases à Asparagines (N) (Rawlings *et col.*, 2011). La dénomination utilisant la lettre P, parfois utilisée, comprend des peptidases utilisant une combinaison de plusieurs types de résidus (Rawlings *et col.*, 2014).



## Chapitre I - Introduction

Les clans et familles sont formés de groupes de peptidases homologues. Les différents clans sont définis essentiellement sur la base de similarités structurales, parfois de séquences ou d'arrangement du site actif. Les noms des clans sont formés d'une lettre majuscule définissant le type catalytique puis d'une deuxième lettre majuscule définissant le clan. Chaque clan est constitué d'une ou plusieurs familles de peptidases dérivant d'une même peptidase originelle. Les familles de peptidases sont définies par homologie de séquences à une peptidase référence, ou holotype, représentant la famille. Les noms des familles sont formés d'une lettre majuscule définissant le type catalytique suivi d'un nombre identifiant la famille. Ces similarités structurales ou de séquences sont basées uniquement sur le domaine de la protéine portant l'activité peptidase (Rawlings *et col.*, 2014).

Pour montrer un exemple, le Tableau 1-1 décrit la hiérarchisation de la famille de peptidases à laquelle appartiennent les peptidases de la famille TET selon MEROPS, en particulier la protéine TET2 de *Pyrococcus horikoshii*.

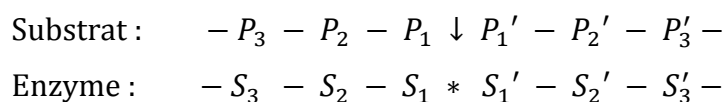
CLAN	FAMILLE	ESPECE	HOLOTYPE
→Clan MH	Famille M18	Espèce M42.001	Glutamyl aminopeptidase
	Famille M20	Espèce M42.002	Aminopeptidase I
	Famille M28	Espèce M42.003	PhTET1 aminopeptidase
	→Famille M42	→Espèce M42.004	PhTET2 aminopeptidase
		Espèce M42.005	TET aminopeptidase
		Espèce M42.006	FrvX putative peptidase
		Espèce M42.007	TM1048-type putative peptidase
		Espèce M42.008	YpdE-type putative peptidase
		Espèce M42.009	PhTET3 aminopeptidase
		Espèce M42.010	CelM peptidase
		Espèce M42.A01	SgcXprotein
		Espèce M42.A02	YhfE
		Homologue non peptidasique	--
		Peptidases non assignées	--

Tableau 1-1 : Classification hiérarchique de la peptidase TET2 de *Pyrococcus horikoshii*. Le clan et la famille auxquels appartiennent TET2 sont indiqués par une flèche.

### 1.2.1 Nomenclature des sites actifs

Afin d'indiquer la spécificité d'une peptidase, le symbole '↓' est utilisé pour indiquer la liaison hydrolysée. La nomenclature utilisée pour décrire la spécificité des peptidases est basée sur un modèle dans lequel on considère la présence de sites spécifiques de part et d'autre du site actif capables d'accueillir la chaîne latérale d'un acide aminé (Berger *et col.*, 1970).

Les sites spécifiques sont annotés, depuis le site actif, S1 ... Sn en direction de l'extrémité N-terminale et S1' ... Sn' en direction de l'extrémité C-terminale. Les résidus accueillis par ces sites sont annotés P1...Pn et P1' ... Pn' respectivement, on peut alors représenter un site actif comme suit : (le site actif est représenté par \*)



## 1.3 PEPTIDASES OLIGOMERIQUES AUTO-COMPARTIMENTEES ATP-INDEPENDANTES

---

Le travail de thèse porte sur les relations structure-fonction et les mécanismes d'assemblage de TET, une peptidase géante auto-compartimentée. Ce chapitre a pour but de présenter l'état des connaissances sur les 8 autres peptidases cytosoliques capables de s'auto-assembler pour former des machines cellulaires. Ces informations permettent de mieux comprendre l'originalité des mécanismes d'activation enzymatique et d'auto assemblage mis à jour dans le cadre de la thèse à travers l'étude du système TET.

### 1.3.1 Peptidase Tricorn

Lors d'une étude visant à identifier des partenaires du protéasome, la peptidase Tricorn a été découverte chez l'Archée *Thermoplasma acidophilum*, elle est la première protéase dégradant les produits du protéasome à avoir été identifiée (Tamura *et col.*, 1996). Sans être un partenaire direct du protéasome, elle participe à la dégradation de ses produits en di- ou tri-peptides. Ces peptides sont ensuite hydrolysés en acides aminés libres par trois peptidases associées à Tricorn, la proline iminopeptidase F1 et les deux metallopeptidases F2 et F3 (Tamura *et col.*, 1998),(Kim *et col.*, 2002) (Figure 1-1).

La peptidase Tricorn est un hexamère de 720 kDa formant une structure rappelant la forme du chapeau tricorne dont elle tire son nom (Figure 1-2B).

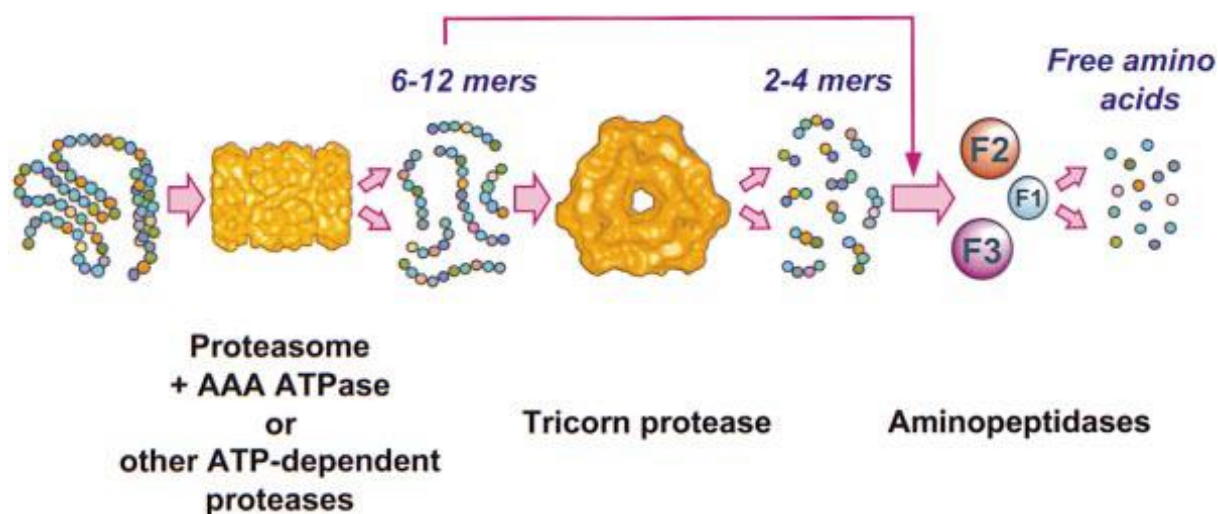


Figure 1-1 : Représentation schématique de la dégradation séquentielle des protéines chez *Thermoplasma acidophilum*. (Tirée de Tamura *et col.* (1998))

Chacune des 6 chaînes polypeptidiques de 120 kDa est constituée de 5 domaines : un  $\beta$ -propeller à six feuillets  $\beta$  ( $\beta_6$ ), un deuxième  $\beta$ -propeller à sept feuillets  $\beta$  ( $\beta_7$ ), un domaine de type PDZ et les domaines C1 et C2 portant les résidus du site actif. (Figure 1-2A) (Brandstetter *et col.*, 2001). On trouve les domaines PDZ dans de nombreuses protéines, il est généralement impliqué dans les interactions protéines-protéines (Ye *et col.*, 2013). La structure cristallographique du Tricorn dévoile deux anneaux trimériques, décalés, s'imbriquant l'un dans l'autre ; chaque sous-unité d'un anneau étant en contact quasi exclusif avec deux sous-unités du deuxième anneau (Figure 1-2B). Les dimères formant le côté long de l'hexamère sont intimement entrelacés. En revanche, l'interface entre les sous-unités formant le côté court de l'hexamère est très plate. En conséquence, les surfaces de ces deux interfaces sont très différentes (6.23 et 1.68 Å), il a donc été proposé que le mode d'assemblage de Tricorn soit un trimère de dimère (Brandstetter *et col.*, 2001). L'hexamère mesure 160 Å selon le plan perpendiculaire à l'axe 3 et 88 Å dans le plan parallèle à ce dernier.

La peptidase Tricorn est également capable de former une superstructure de 14,6 MDa composée de 20 hexamères formant une capsidie icosaédrique de  $55 \pm 2$  nm (Walz *et col.*, 1999). Cette structure existe *in vivo* et a été analysée par microscopie électronique (Figure 1-3)

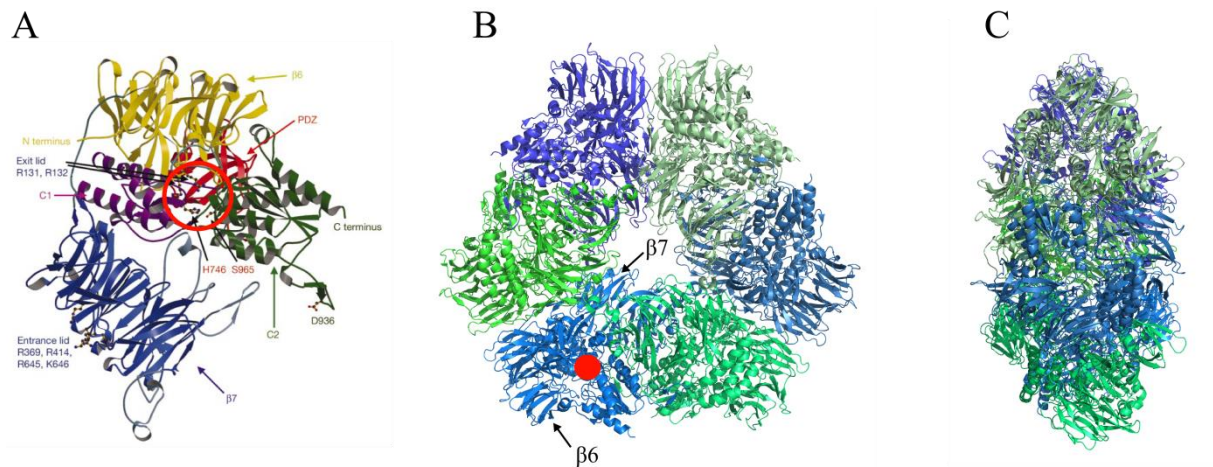


Figure 1-2 : Structure de la peptidase Tricorn. A) Structure monomérique de la Tricorn de *T. acidophilum*, chaque domaine est représenté dans une couleur différente, le site actif est entouré en rouge (code PDB = 1K32). B) Représentation en ruban de la structure de la peptidase Tricorn suivant l'axe 3. La position des  $\beta$ -propeller 6 et 7 sont indiqués par des flèches sur une des sous-unités. La position d'un site actif est indiquée par un point rouge. C) Représentation en ruban de la structure de la peptidase Tricorn vue sur la tranche. (Tirée de Brandstetter *et col.* (2001))

Tricorn est une endopeptidase à sérine appartenant à la famille S41 selon la classification MEROPS (Rawlings *et col.*, 2014). Elle présente une activité de type trypsine prépondérante, mais est aussi capable de cliver des substrats de type chymotrypsine (Tamura *et col.*, 1996). Le site actif, relativement enfoui dans la structure, est situé à l'interface entre les domaines C1 et C2 et est formé d'une tétrade catalytique : Ser745, His746, Ser965 et Glu1023.

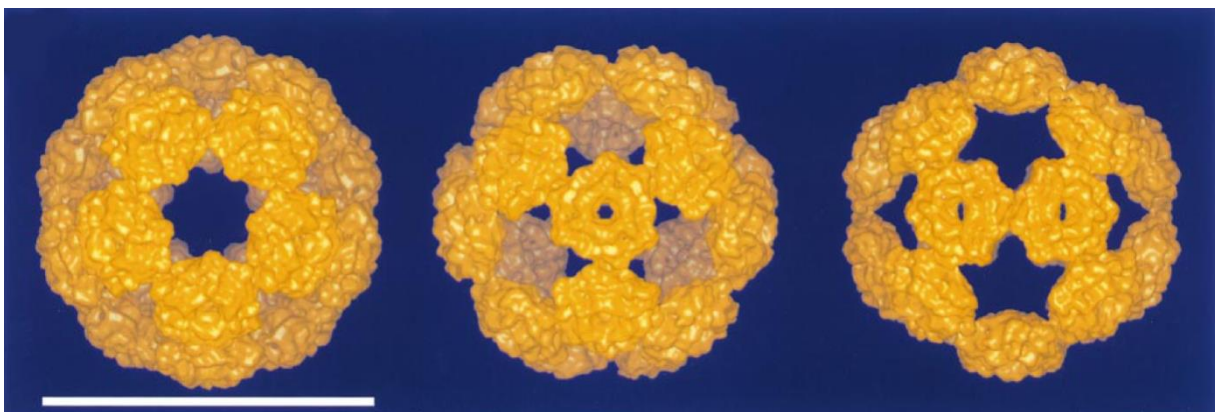


Figure 1-3 : Représentation de l'iso surface de la capsid de Tricorn à une résolution de 1.3 nm. Vues suivant les axes de symétrie 5 (gauche), 3 (centre) et 2 (droite). La barre d'échelle correspond à 50 nm. (Tirée de Walz *et col.* (1999))

Les structures de mutants et de complexes avec des inhibiteurs ont permis de proposer un mécanisme d'entrée des substrats et de sortie des produits. Les axes des  $\beta$ -propeller

$\beta 7$  et  $\beta 6$  étant dirigés vers le site actif, il a été proposé que la sélection et l'accès du substrat au site actif se ferait par le biais du  $\beta$ -propeller  $\beta 7$ , le produit serait ensuite libéré par le  $\beta$ -propeller  $\beta 6$  (Kim *et col.*, 2002). Les activités enzymatiques de l'hexamère et de la capside de Tricorn sont identiques, le rôle de la capside reste quant à lui indéterminé.

Les protéines F1, F2 et F3 qui ont été proposées comme partenaires du Tricorn sont des peptidases monomériques possédant chacune une spécificité de substrat différente. La protéine F1 est une proline aminopeptidase à sérine de 33.5 kDa (famille S33). Les protéines F2 et F3 sont des métallo-aminopeptidases homologues d'environ 89 kDa (famille M1). F2 possède une spécificité de substrat large puisqu'elle est active contre les résidus neutre, hydrophobe et basique. F3 présente quant à elle une spécificité de substrat moins large puisqu'elle n'est active que contre les résidus acides (Tamura *et col.*, 1998).

La peptidase Tricorn ne semble être présente que chez certaines Archées thermophiles. On retrouve des séquences orthologues à celle de *Thermoplasma acidophilum* chez les organismes de la classe des *thermoplasmata* et dans plusieurs organismes de la classe des *thermoprotei*, notamment chez tous les organismes appartenant au genre *sulfolobus* ainsi que chez quelques *pyrobaculum*. Les peptidases F1, F2 et F3 ont quant à elles des homologues dans les 3 domaines du vivant. Le rôle physiologique du Tricorn reste à ce jour indéterminé.

### 1.3.2 Tripeptidyl peptidase II

La tripeptidyl peptidase II ou TPPII fait partie des peptidases prenant en charge les produits de dégradation du protéasome chez les Eucaryotes (Tomkinson, 1999). C'est une peptidase à sérine appartenant à la famille S8, aussi appelée famille des subtilases, du clan SB (MEROPS). C'est une aminopeptidase possédant une activité exopeptidase qui lui permet de cliver des tripeptides à l'extrémité N-terminale de peptides issus du protéasome, mais elle possède également une activité endopeptidase, cette dernière reste aujourd'hui peu caractérisée (Geier *et col.*, 1999). C'est l'unique exopeptidase appartenant à la famille des subtilases.



C'est une grande aminopeptidase de 128 à 150 kDa, dépendant de l'espèce. Chaque monomère est constitué de 3 domaines. Le domaine N-terminal correspondant au domaine subtilisine entrecoupé d'une insertion forme une structure en anneau dans lequel est situé le site actif. Les deux autres domaines sont le domaine central et le domaine C-terminal (Rockel *et col.*, 2012) (Figure 1-4).

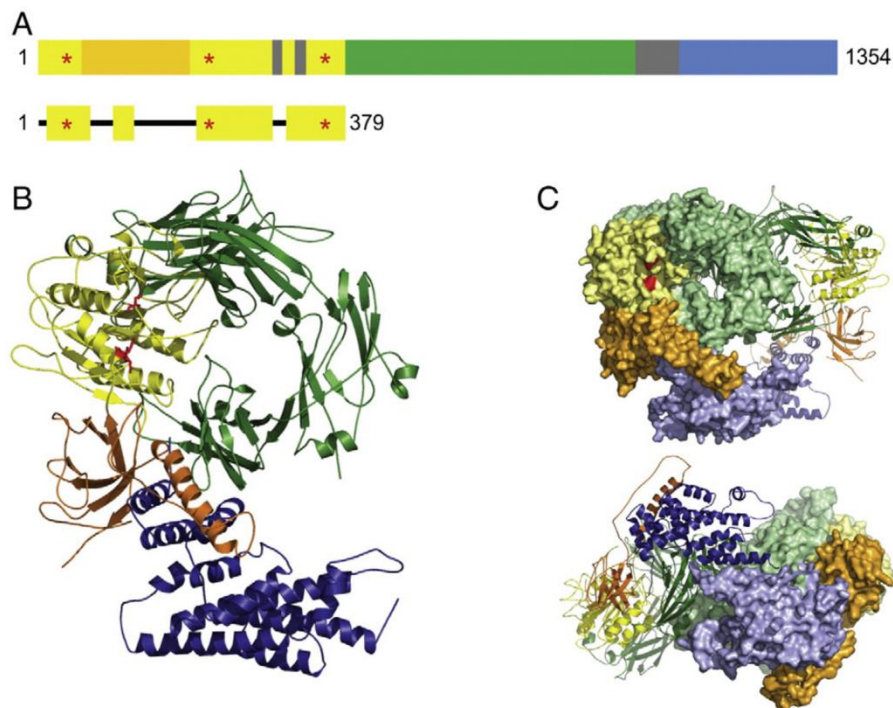


Figure 1-4 : Structure de la peptidase TPPII (Code PDB : 3LXU) : A) Composition des domaines constituant TPPII comparé à la subtilisine. Première barre : Jaune : domaine de type subtilisine de TPPII, les acides aminés du site actif sont représentés par des astérisques ; orange : insertion dans le domaine du site actif ; vert : domaine central ; bleu : domaine C-terminal ; gris : boucle absente de la structure cristallographique de TPPI (code PDB : 3LXU). Deuxième barre : jaune : domaine de la subtilisine Carlsberg aligné aux domaines de TPPII, les acides aminés du site actif sont représentés par des astérisques. B) Représentation en ruban de la structure de TPPII de *Drosophila melanogaster*, les différentes couleurs des domaines sont décrites en A). C) Représentation du dimère de TPPII de *Drosophila melanogaster*, les différentes couleurs des domaines sont décrites en A). Un des monomères est représenté en ruban, le second en surface. (Tirée de Rockel *et col.* (2012))

Dans la cellule, les sous-unités de TPPII s'assemblent pour former un grand complexe homo-oligomérique dont le poids moléculaire est d'environ 6 MDa. L'enveloppe de ce complexe a été résolue par microscopie électronique et révèle une structure originale formant une double hélice. Ce complexe est constitué de 2 brins hélicoïdaux de 60 nm de long, chaque brin étant constitué de 10 dimères de TPPII empilés. La structure cristallographique du dimère de TPPII a pu être résolue et « dockée » dans l'enveloppe obtenue par microscopie électronique (Chuang *et col.*, 2010) (Figure 1-5A). Cette structure hybride a permis de mettre en évidence un réseau complexe de cavités

renfermant les sites actifs de TPPII au sein de chacun des deux brins du complexe, ceci permettant de protéger le protéome de la cellule (Rockel *et col.*, 2012) (Figure 1-5B).

TPPII n'est active que sous sa forme oligomérique, le dimère de TPPII ne présente quant à lui que 10 % de l'activité du complexe. C'est l'oligomérisation de la particule et la formation du réseau de cavités internes qui permet l'activation de la peptidase TPPII. L'oligomérisation permet notamment la formation des poches catalytiques, la formation du réseau de canaux internes permettant la reconnaissance et le guidage des peptides substrats, la réorientation d'une boucle interne et enfin le déplacement de la Ser462 du site actif en position optimale pour l'activité de la peptidase (Figure 1-5C).

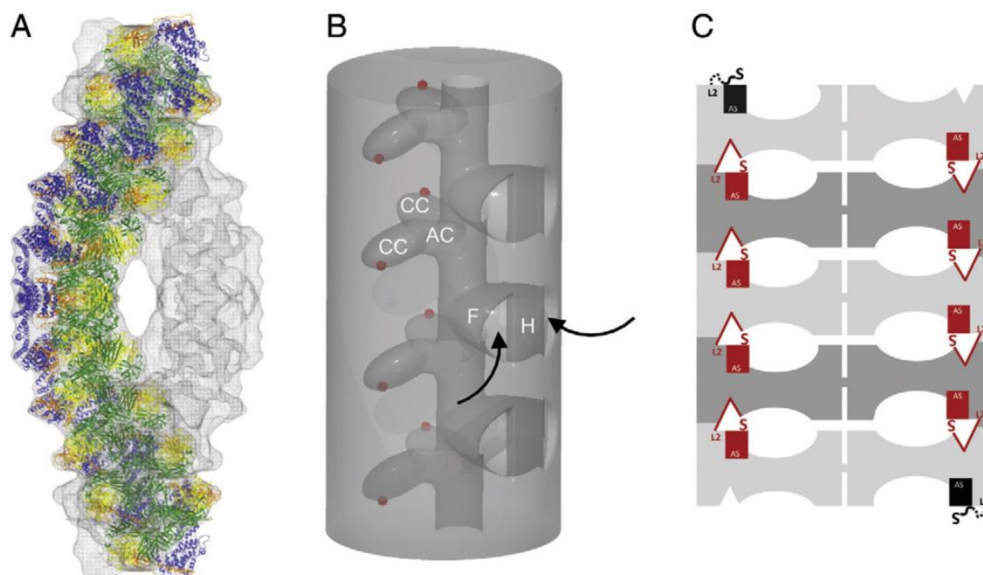


Figure 1-5 : Structure hybride de l'hélice de TPPII. A) Structure cristallographique du dimère de TPPII de *Drosophila melanogaster* (représenté en ruban) « dockée » dans l'enveloppe de l'hélice de TPPII (représentée en maille) obtenue par microscopie électronique. Les couleurs correspondent à celles de la figure 4. B) Représentation schématique du réseau de cavité formé lors de l'assemblage des dimères de TPPII en brins. Points rouges : sites actifs ; H : poignée ; CC : poches catalytiques ; AC : antichambre ; F : foyer. Les flèches indiquent les entrées du réseau de cavité de chaque côté des poignées du dimère. C) Représentation schématique de l'activation de TPPII : l'empilement des dimères mène à la formation du réseau de cavités, à la réorientation d'une boucle ainsi qu'au placement correct du résidu Ser462 du site actif. (Tirée de Rockel *et col.* (2012))

Par ailleurs il est intéressant de noter que l'insertion du domaine supplémentaire au niveau du site actif (Figure 1-4A) pourrait jouer un rôle déterminant dans la formation du complexe TPPII (Tomkinson, 1999), nous verrons que c'est aussi le cas pour les TET. Son rôle physiologique n'a pas été déterminé.

### 1.3.3 DegP / HtrA

La peptidase DegP (degradation of periplasmic proteins), aussi nommé HtrA (high temperature requirement A), est une peptidase à sérine de la sous-famille S1C du clan PA(S) (MEROPS). Elle est découverte chez *E. coli* à la fin des années 80 lors de l'étude de souche d'*E. coli* dont la dégradation des protéines périplasmiques était déficiente ou qui était incapable de croître à température élevée (Schuhmann *et col.*, 2012), elle est le premier membre de la famille des protéines HtrA à avoir été découverte.

Elle est conservée dans les 3 domaines du vivant. C'est une endopeptidase périplasmique essentielle à *E. coli* pour croître à haute température (Lipinska *et col.*, 1990). Cette peptidase présente la particularité de posséder une activité chaperonne, en plus de son activité peptidase. La bascule entre les deux activités est dépendante de la température (Spiess *et col.*, 1999). En effet, en dessous de 28 °C, l'activité peptidase est très faible et l'activité chaperonne de DegP est maximale ; au-delà de 32 °C, l'activité endopeptidique de DegP augmente et devient prépondérante alors que l'activité chaperonne diminue.

Le monomère de la peptidase DegP est composé d'un domaine N-terminal dans lequel se trouve le site actif portant l'activité de type chymotrypsine, la partie C-terminal de la protéine est formée de deux domaines PDZ (PDZ1 et PDZ2). Une boucle flexible connecte le domaine catalytique à PDZ1 et une seconde connecte PDZ1 à PDZ2 (Figure 1-6A).

En l'absence de substrat, la peptidase DegP forme un homo-hexamère composé de deux anneaux trimériques (Krojer *et col.*, 2008). Cet hexamère est une forme inactive de la peptidase dans laquelle le site actif n'est pas structuré correctement et où une des boucles flexibles masque le site actif. La fixation d'une protéine dénaturée ou mal repliée sur les domaines PDZ provoque l'activation de la peptidase. Lors de cette activation, l'assemblage de la peptidase DegP est modifié. La forme active de la peptidase a la forme d'une cage polyédrique formée de 4 ou 8 trimères de DegP. Le complexe à 24 sous-unités d'environ 1,3 MDa ressemble à une grande sphère dont le volume de la cavité centrale est d'environ 700 000 Å<sup>3</sup>.



Les monomères de DegP s'assemblent sous forme de trimères via des réseaux d'interaction hydrophobes, le trimère semble être la brique de base pour la formation des autres complexes (Figure 1-6B) (Krojer *et col.*, 2002). Au sein de l'hexamère, les contacts entre les deux trimères se font essentiellement via des longs feuillets  $\beta$  mobiles issus du domaine peptidase, interagissant deux par deux, ils s'ancrent dans le site actif de la sous-unité de l'anneau opposé (Figure 1-6). Entrent également en jeu des interactions PDZ1-PDZ1 et PDZ2-PDZ2 entre les deux anneaux trimériques, cependant ces domaines restent très mobiles. Les deux trimères étant fixés par des structures mobiles, il en résulte un complexe relativement lâche (Krojer *et col.*, 2002).

Pour comprendre comment se forment les complexes à 12 ou 24 sous-unités, la structure du monomère a été adaptée dans l'enveloppe des complexes obtenue par microscopie électronique. Cette étude révèle que la structure des trimères et les contacts nécessaires à la formation des complexes 12-mère et 24-mère sont quasi-identiques (Jiang *et col.*, 2008).

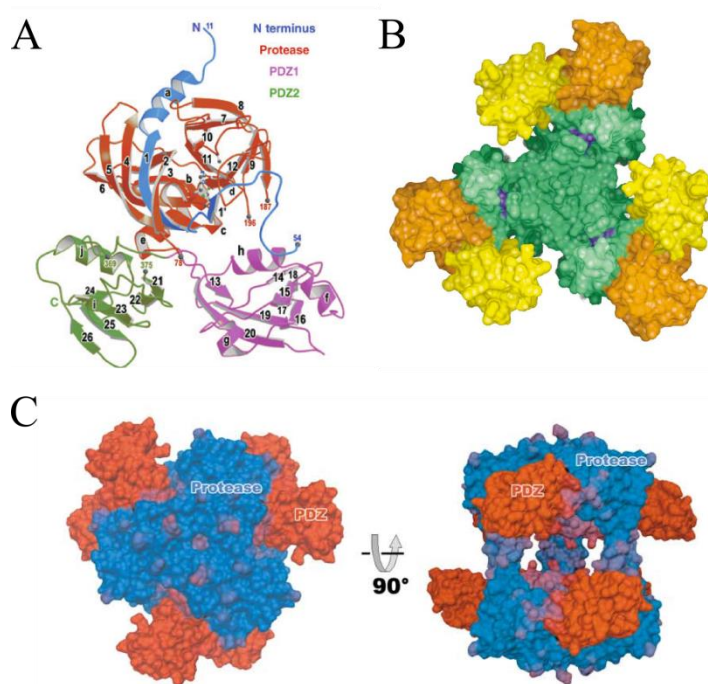


Figure 1-6 : Structure et assemblages des formes inactives de la peptidase DegP (Code PDB: 1KY9). A) Structure du monomère de DegP représenté en ruban. Les couleurs correspondent aux différents domaines, les structures secondaires sont indiquées. (tirée de Groll *et col.* (2005)). B) Structure du trimère représenté avec l'enveloppe moléculaire. La localisation des différents domaines (catalytique en vert, PDZ1 en jaune, PDZ2 en orange) ainsi que les sites catalytiques (bleu) sont indiqués (tirée de (Clausen *et col.*, 2002)). C) Structure de l'hexamère représentée avec l'enveloppe moléculaire coloré en fonction du thermal motion factor (bleu : région rigide, rouge : région flexible). (Tirée de Clausen *et col.* (2002))

Dans les deux cas, ce sont les interactions entre les domaines PDZ1 d'un trimère et PDZ2 d'un second trimère qui structurent le complexe (Figure 1-7). L'interaction PDZ1-PDZ2 est rendue possible grâce à la réorientation de ces domaines qui, relativement flexibles

dans l'hexamère, deviennent ici très rigides à cause de la torsion nécessaire à leur interaction, rendant ainsi les structures à 12 ou 24 sous-unités plus rigides que l'hexamère (Jiang *et col.*, 2008).

Lors de l'assemblage des complexes à 24 sous-unités, les sites actifs ne sont accessibles que depuis l'intérieur de la particule. De larges pores de 35 Å ou 14 Å sont formés permettant l'entrée des substrats dans la particule, délimités par les domaines PDZ1 et PDZ2. Les domaines PDZ1 reconnaissent et fixent les extrémités C-terminales exposés de protéines dénaturées ou mal repliées. Par effet allostérique, la fixation d'un substrat sur le domaine PDZ1 provoque l'activation du site actif associé. Enfin, le substrat « tenu » par le domaine PDZ1 peut atteindre le site actif et être clivé par celui-ci. La peptidase DegP clive préférentiellement les résidus valine, alanine, thréonine et isoleucine, ceux-ci font partie des acides aminés les plus répandus dans les protéines. Le produit ainsi relâché à l'intérieur de la particule peut être fixé par un autre domaine PDZ1 et son clivage se poursuivre. DegP est donc une peptidase non-processive (Krojer *et col.*, 2008).

DegP ou HtrA fait partie de la famille des protéines HtrA, à laquelle elle a d'ailleurs donné son nom. DegP est l'homologue bactérien de cette famille de protéine et a été identifié et caractérisé chez *E. coli*. Chez cette dernière, il existe deux homologues de DegP, qui sont DegS et DegQ. DegS est ancré à la face interne du périplasme et agit comme senseur pour alerter la cellule en cas d'accumulation de protéine mal repliée. DegQ a une structure et une activité quasi identique à celle de DegP, il a été proposé qu'elles puissent avoir exactement le même rôle. Chez l'homme, l'homologue de DegP est nommé HtrA1 (Clausen *et col.*, 2002; Krojer *et col.*, 2002). Chez les plantes, il a été montré que les formes Deg5 et Deg8 peuvent s'assembler sous la forme d'un hétéro-oligomère dans le lumen des thylakoïdes, il a été proposé que ce complexe joue un rôle dans la réparation du photo-système II (Sun *et col.*, 2007).

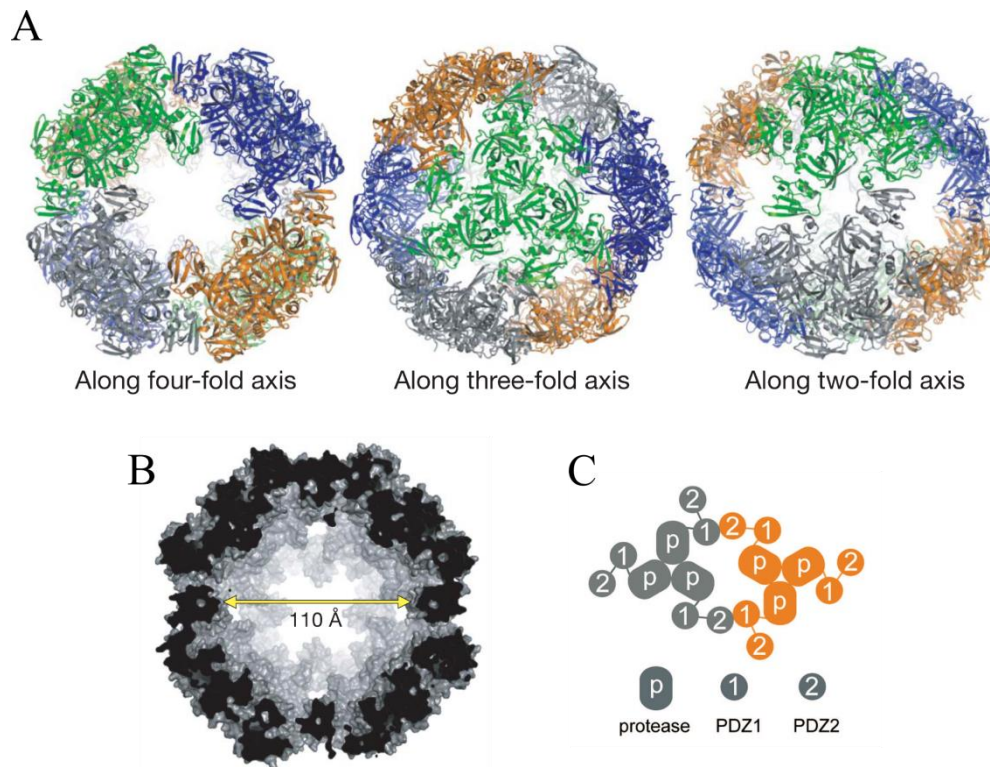


Figure 1-7: Détails de la structure du complexe à 24 sous-unités de DegP (Code PDB: 3CS0). A) Représentation en ruban de la structure du complexe à 24 sous-unités de DegP. Chaque unité trimérique est colorée différemment. La particule est représentée selon trois orientations différentes (gauche : axe de symétrie 4, milieu : axe de symétrie 3, droite : axe de symétrie 2). B) Représentation de la capsule à 24 sous-unités en coupe transversale. Le diamètre de la cavité centrale est indiqué par une flèche jaune. C) Représentation schématique des interactions intra-trimère entre les domaines catalytiques et inter-trimère entre les domaines PDZ1 et PDZ2 dans les complexes à 12 ou 24 sous-unités. (Tirée de Krojer *et col.* (2008) et Jiang *et col.* (2008))

#### 1.3.4 Bloemycin hydrolase / Gal6

La bleomycine hydrolase est une aminopeptidase à cystéine appartenant à la famille C1 du clan CA, elle fait partie du même clan que la papaïne dont le site actif est très similaire. Elle tient son nom de sa capacité à hydrolyser, et ainsi désactiver, un glycopeptide antibiotique, la bleomycine (O'Farrell *et col.*, 1999). L'homologue de la bleomycin hydrolase chez la levure est la peptidase Gal6, étonnamment, elle a été purifiée en tant que protéine liée à l'ADN au niveau de la séquence d'activation en amont du system GAL. Lorsqu'elle est surexprimée chez la levure, elle lui confère une résistance accrue à la bleomycine (Kambouris *et col.*, 1992). La caractéristique originale de la forme exprimée chez la levure est donc sa capacité à fixer des acides nucléiques (ADN simple brin, ARN simple brin et ADN double brin clivé) (Joshua-Tor *et col.*, 1995).

Le monomère de Gal6 est formé d'un domaine catalytique présentant une similarité structurale avec la papaïne, un domaine d'oligomérisation, un domaine comprenant des hélices alpha et le domaine C-terminal ayant la forme d'un bras (Figure 1-8A) (Joshua-Tor *et col.*, 1995).

La bleomycine hydrolase s'assemble sous la forme d'un homo-hexamère, qui peut être considéré comme un trimère de dimère. L'oligomérisation permet la formation d'un large pore central donnant accès à une cavité d'où les sites actifs sont accessibles (Figure 1-8B, Figure 1-9A). Lors de la formation du complexe, le domaine d'oligomérisation d'un monomère se fixe en s'enroulant au domaine d'hélices alpha d'un second monomère et vice versa, créant ainsi une interaction très forte entre les monomères d'un dimère (Figure 1-8B et C). Les dimères sont ordonnés autour d'un axe de symétrie 3 pour former l'hexamère de Gal6, les contacts entre dimères font intervenir des interactions de type coiled-coil (Figure 1-8) (Joshua-Tor *et col.*, 1995).

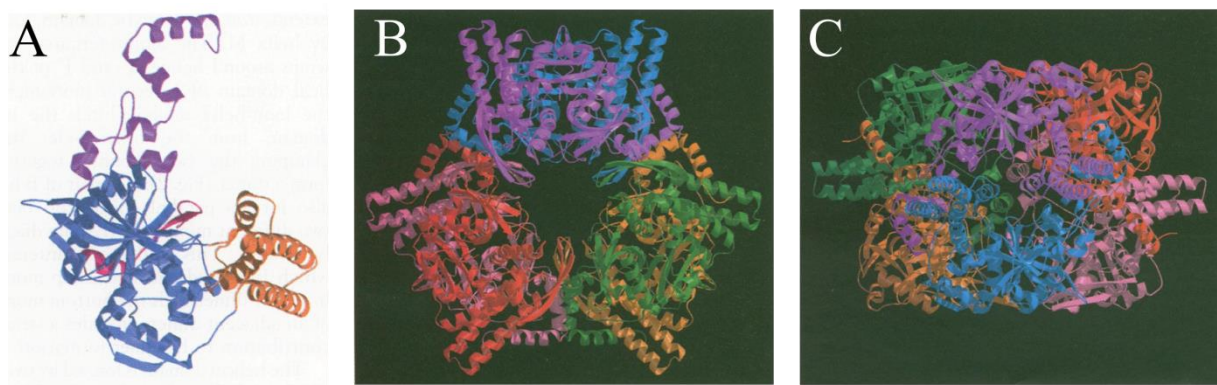
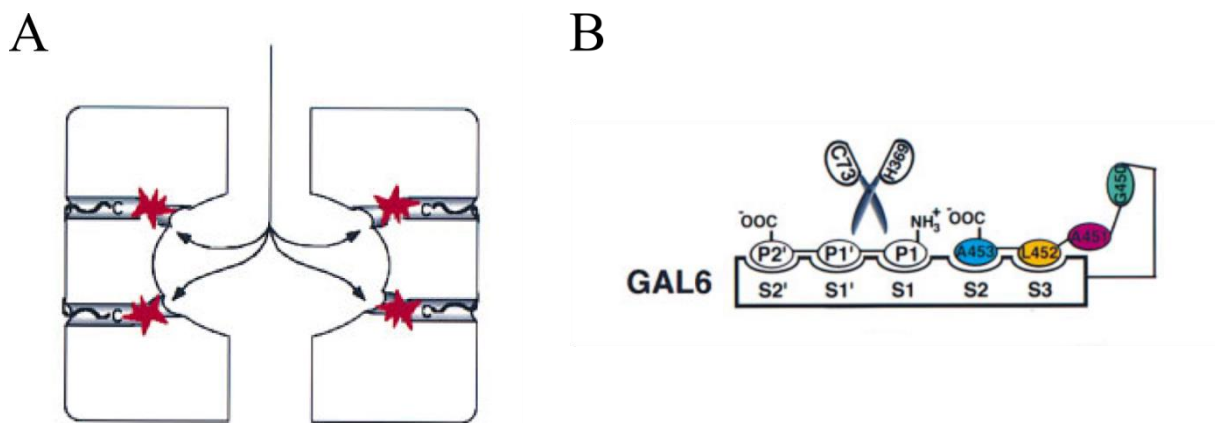


Figure 1-8 : Structure cristallographique de la peptidase Gal6 (Code PDB : 1GCB). A) Structure du monomère de Gal6 représentée en ruban. Le domaine d'oligomérisation est représenté en pourpre, le domaine catalytique en bleu, le domaine d'hélice alpha en jaune et le bras C-terminal en rouge. B) Structure de l'hexamère de Gal6 représenté en ruban vu selon l'axe 3 de symétrie. Le canal central est visible. Les trois dimères formant l'anneau sont respectivement représentés en bleu et pourpre, rouge et rose, vert et jaune. Les interfaces entre les deux anneaux trimériques sont situés entre deux monomères positionnés en diagonale, ils sont situés entre les sous-unités rouge et bleu, rose et verte, jaune et pourpre. C) Vue perpendiculaire à B). Les interactions entre dimères sont ici visibles. On peut voir le domaine d'oligomérisation de la sous-unité pourpre s'enrouler autour du domaine d'hélice alpha de la sous-unité bleue. (Tirée de Joshua-Tor *et col.* (1995))

Les particularités inhabituelles de cette peptidase sont son activation et son mode d'action liés à la présence de l'extrémité C-terminale au niveau du site actif. En effet, l'extrémité C-ter de la bleomycine hydrolase Gal6 s'immisce dans le site actif et peut s'y fixer (Figure 1-9B). Le site actif du précurseur Gal6p est correctement replié et accessible au substrat et présente une activité endopeptidase. Cependant, lors de sa

maturation, la peptidase va agir comme une carboxypeptidase sur sa propre extrémité C-terminale. Cette action aura pour effet de fixer l'extrémité C-ter dans le site actif ou elle occupe les sites S2 à S4 de fixation du substrat. Dans ce cas particulier, l'habituelle activité endopeptidase du domaine catalytique de type papaïne devient une activité aminopeptidase. En effet, par la suite, l'extrémité C-terminale, chargée négativement, ancre l'extrémité N-terminale, chargée positivement, du peptide substrat dans le site actif qui sera clivé. Il a été proposé que le site de clivage du peptide puisse être médié par la position de l'extrémité C-terminale par rapport aux résidus du site actif ; cette conformation pourrait permettre de modifier la taille des produits issus de la catalyse du peptide substrat (Zheng *et col.*, 1998).



**Figure 1-9 : Représentation schématique des sites actifs de Gal6. A) Représentation schématique d'une coupe transversale de l'hexamère de Gal6. Quatre des six sites actifs sont représenté par des étoiles et l'extrémité C-ter s'immisçant dans le site actif est représentée. L'accès aux sites actifs se fait par le pore central. B) Représentation schématique du site actif de Gal6. L'extrémité C-ter se fixant au site actif est représenté en couleur. Les ciseaux représentent le site actif porté par les résidus Cys36 et His369, les ovales non colorés représentent les résidus P1, P1' et P2' du peptide substrat. (Tirée de Zheng *et col.* (1998))**

Le mode d'action présenté ici est celui caractérisé chez Gal6, il est identique chez la forme humaine de la bleomycine hydrolase (O'Farrell *et col.*, 1999). La peptidase bleomycine hydrolase est conservée dans l'ensemble du domaine Eucaryote et a été détectée chez certaines bactéries de ferment lactique (Chapot-Chartier *et col.*, 1994). Tous les homologues de cette famille de peptidases n'ont cependant pas conservé l'activité de liaison à l'ADN, la bleomycine hydrolase humaine a perdu cette capacité (O'Farrell *et col.*, 2007).

Le rôle physiologique de la bleomycine hydrolase est resté obscur pendant plusieurs années. Cependant, il a été récemment démontré que cette peptidase pourrait jouer des



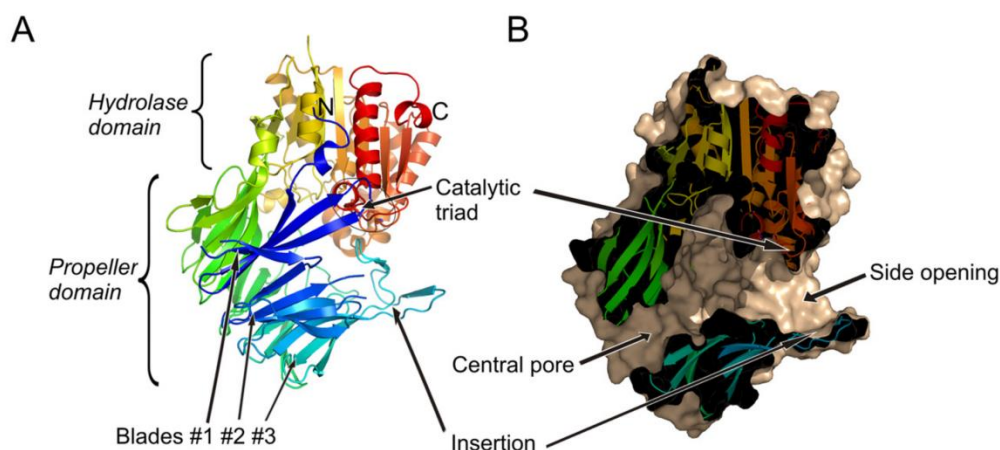
rôles importants dans la régulation de la maturation de précurseurs de protéines amyloïdes, dans la maturation post-traductionnelle de peptides en vue de leur présentation au complexe majeur d'histocompatibilité I, dans les mécanismes de protection contre la toxicité de l'homeocystéine, dans la production d'acides aminés libres agissant comme agents d'hydratation et dans la production de fragment N-terminaux court d'huntingtine (Montoya *et col.*, 1998; Namba *et col.*, 1999; Lefterov *et col.*, 2001; Zimny *et col.*, 2006; Towne *et col.*, 2007; Kim *et col.*, 2009; Ratovitski *et col.*, 2011).

### 1.3.5 Trilobed peptidase / PhAAP

La trilobed peptidase, aussi appelé PhAAP, est une peptidase à sérine faisant partie de la famille des prolyl-oligopeptidases (POP) de la famille S9 du clan SC. Elle est aussi nommée PhAAP pour *Pyrococcus horikoshii* acylaminoacyl peptidase. Il existe deux peptidases dénommées trilobed peptidase chez *Pyrococcus horikoshii* dans la base de données MEROPS. Une première, codée par le gène PH0863, a été caractérisée par Ishikawa *et col.* (1998) et est active sous la forme d'un dimère. La deuxième, codée par le gène PH0594, est quant à elle active sous la forme d'un hexamère ayant une structure trilobée dont elle tire son nom. La première peptidase caractérisée est bien une peptidase de la famille des acylaminoacyl peptidases, mais elle a été dénommée par erreur trilobed peptidase à cause de sa similarité avec la seconde.

Le monomère de la peptidase trilobed est constitué de deux domaines : un domaine  $\beta$ -propeller à 7 feuillets  $\beta$  et un domaine catalytique en C-terminal. Une large cavité d'environ 22 Å de diamètre sépare les deux domaines, elle contient le site actif (Figure 1-10). En plus du pore formé par le domaine  $\beta$ -propeller, dénommé central pore sur la figure 1-10, un deuxième pore d'environ 12 Å est présent entre les domaines  $\beta$ -propeller et catalytique, dénommé side opening sur la figure 1-10 (Menyhard *et col.*, 2013).

La structure cristallographique de PhAAP révèle un complexe hexamérique formé d'un trimère de dimère de PhAAP (Figure 1-11A et B). L'analyse des interfaces de contacts a montré que les interactions entre les monomères au sein du dimère sont plus fortes que celles responsables de la formation du trimère.

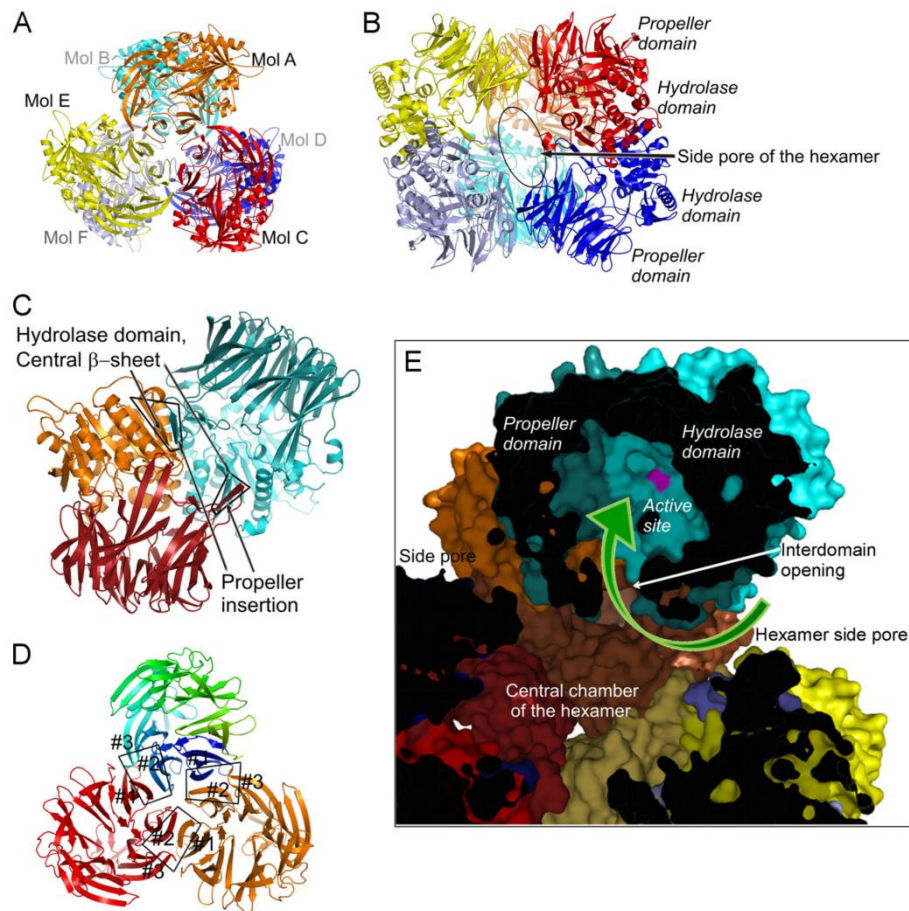


**Figure 1-10 : Structure du monomère de PhAAP (Code PDB : 4HXE). A) Représentation en ruban. B) Coupe transversale du monomère et représentation de la surface moléculaire. Les deux pores liés à la cavité interne du monomère sont représentés (un pore porté par le domaine  $\beta$ -propeller et un second formé entre le domaine catalytique et les feuillets #1 et #3 du domaine  $\beta$ -propeller. (Tirée de Menyhard *et col.* (2013))**

L'interface entre les dimères est formée par la longue boucle du feuillet #3 du domaine  $\beta$ -propeller d'un monomère et le domaine catalytique d'un second monomère Figure 1-11C. L'interface entre les trimères est quant à lui formé par l'interaction entre le feuillet #1 du domaine  $\beta$ -propeller d'un monomère faisant face aux feuillets #2 et #3 d'un second monomère (Figure 1-11D). La formation du trimère, en restreignant les mouvements des feuillets  $\beta$  1-3 du domaine  $\beta$ -propeller, contribue à fixer la taille du pore donnant accès au site actif.

Lors de l'assemblage de l'hexamère, 3 larges pores mesurant entre 20 et 30 Å sont formés (Figure 1-11B et E), ils forment des canaux à l'intérieur de la particule et se joignent pour former une cavité centrale. Chacune des 6 poches catalytiques sont tournées vers l'intérieur de la particule et sont donc connectées à la cavité centrale de l'hexamère. Chaque monomère possède 2 pores différents (Figure 1-10), cependant, la flexibilité du domaine  $\beta$ -propeller étant restreinte par la formation du trimère, celui-ci ne peut pas permettre l'entrée de substrat. Finalement, il a été proposé que le criblage des substrats se fasse donc en deux temps. Les substrats assez petits pénètrent la particule par les larges pores situés à la surface de l'hexamère, ils sont ensuite conduits vers la cavité centrale via les canaux formés, enfin, les peptides assez flexibles pour pénétrer les poches catalytiques sont clivés, les peptides trop gros ou encore repliés peuvent sortir de la particule par un autre canal.

La « trilobed peptidase » possède une activité acylaminoacyl peptidase, elle est donc capable de cliver les résidus des extrémités N-terminales ayant été acylés, et présente une préférence pour les résidus ayant une longue chaîne hydrophobe comme la phénylalanine ou la leucine (Szeltner *et col.*, 2009). Contrairement aux autres membres de sa famille de peptidases, la trilobed peptidase possède aussi une activité endopeptidase (Szeltner *et col.*, 2009).



**Figure 1-11 : Structure de l'hexamère de PhAAP (Code PDB : 4HXE).** A, B) Vue de face et de côté de l'hexamère constitué d'un trimère de dimères. Trois larges pores mènent à la cavité centrale du complexe. C) Le dimère est principalement stabilisé par l'interaction entre le domaine  $\beta$ -propeller d'un monomère et le domaine catalytique d'un second monomère. D) Le trimère est stabilisé par les interactions entre les feuillets 1 à 3 des domaines  $\beta$ -propeller. E) Coupe transversale de l'hexamère représentée en surface moléculaire. La flèche verte représente le cheminement d'un peptide vers la poche catalytique de la peptidase. (Tirée de Menyhard *et col.* (2013))

Enfin, la trilobed peptidase ou PhAAP est pour l'instant le seul exemple de peptidase de la famille des POP à former un complexe hexamérique. La peptidase Eucaryote dipeptidyl-peptidase IV est le seul autre membre à former un grand assemblage, en effet la structure de DPP IV est tétramérique et semble adopter un mode d'action proche de celui de la trilobed peptidase (Menyhard *et col.*, 2013). On retrouve des AAP dans les



trois domaines du vivant, cependant seules les formes bactériennes et archéennes semblent pouvoir former le complexe trilobé. Le rôle de la « trilobed peptidase » reste à ce jour inconnu.

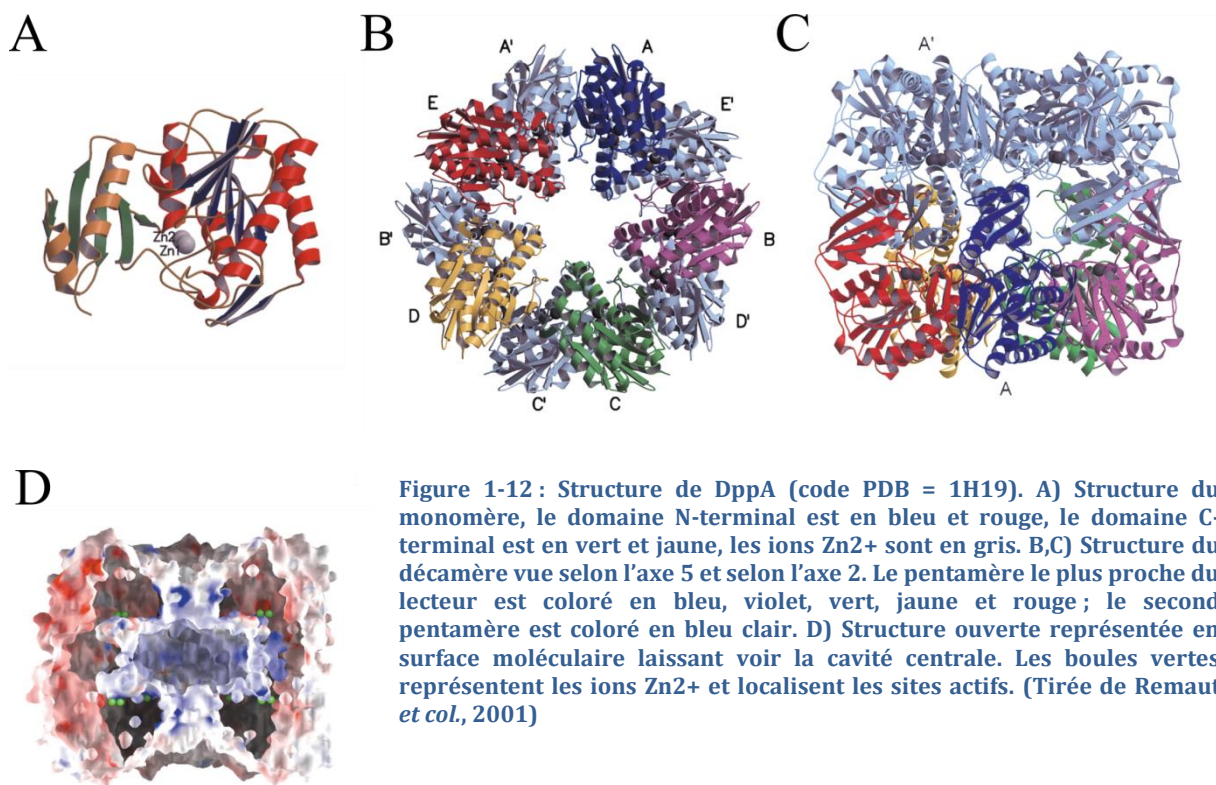
### 1.3.6 DppA

La peptidase DppA est une aminopeptidase découverte chez *Bacillus subtilis*, c'est une metallopeptidase à zinc de la famille M55 du clan MN.

La structure cristallographique résolue de DppA révèle un complexe homo-oligomérique à 10 sous-unités arrangées en tonneau. Chaque monomère est constitué de deux domaines : un domaine N-terminal portant le site actif et les résidus coordonnant les atomes de zinc  $Zn^{2+}$  ; et un domaine C-terminal impliqué dans la formation de la structure quaternaire (Remaut *et col.*, 2001).

Le décimère est formé de deux pentamères reliés par une symétrie d'ordre 2 (Figure 1-12B et C), l'ensemble des contacts entre les sous-unités implique des interactions hydrophobes, électrostatiques ainsi que des liaisons hydrogènes. Le domaine N-terminal est exclusivement responsable des interactions entre sous-unités au sein du pentamère, huit feuillets  $\beta$  d'un monomère interagissant avec deux feuillets  $\beta$  d'un second monomère, et permet ainsi la formation du pentamère en forme d'anneau. Les domaines C-terminaux pointent vers l'extérieur de l'anneau et interagissent avec ceux d'un deuxième anneau par contact hydrophobe et électrostatique pour permettre la formation du décimère (Remaut *et col.*, 2001). La particule ainsi formée mesure 110 x 110 x 85 Å<sup>3</sup>. Un canal de 20 Å de diamètre traverse le complexe suivant l'axe de symétrie 5 et donne accès à la cavité centrale de 50 Å (Figure 1-12D). Les sites actifs sont disposés en zigzag le long de la paroi interne de cette cavité. L'accès aux sites catalytiques de la cavité est donc restreint par les pores du décimère.

La peptidase DppA possède une activité exclusivement D-spécifique et hydrolyse préférentiellement les substrats D-Ala-D-Ala et D-Ala-Gly-Gly (Cheggour *et col.*, 2000). Des séquences homologues ont été identifiées chez d'autres Bactéries ainsi que chez des Archées, mais aucune ne l'a été chez des Eucaryotes. Le rôle physiologique de DppA est à ce jour inconnu.

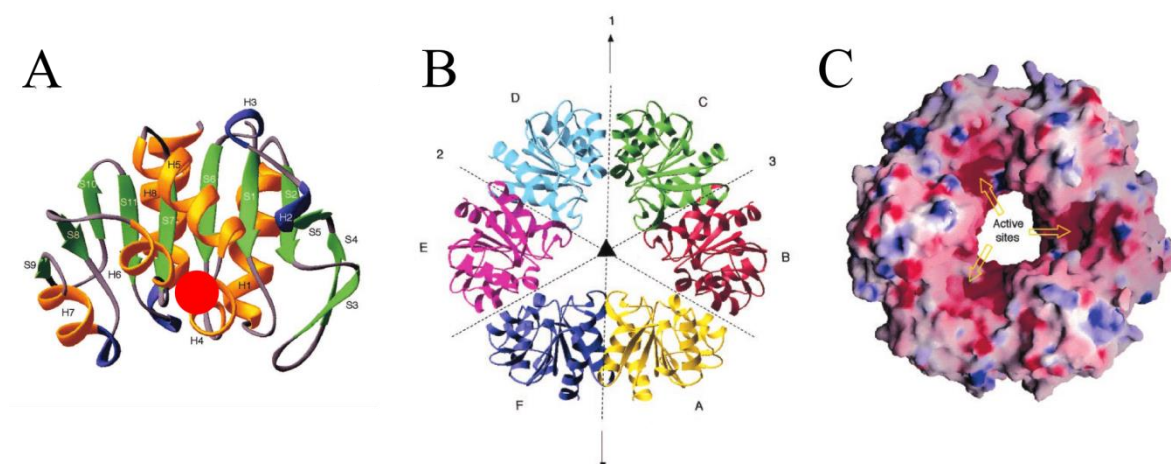


### 1.3.7 PhP1 / PfP1

La peptidase Protease I, aussi dénommée PfP1 ou PhP1 pour respectivement *Pyrococcus furiosus* Protease I et *Pyrococcus horikoshii* Protease I, est une aminopeptidase à sérine de la famille C56 du clan PC, elle est active préférentiellement contre les résidus basiques et hydrophobes (Arg) en position P1. Elle montre également une activité endopeptidase, plus faible, contre les mêmes types de résidus (Zhan *et col.*, 2014).

Le monomère de Protease I de 18.8 kDa est formé d'un sandwich d'hélices  $\alpha$  et de feuillets  $\beta$  en forme de croissant, le site actif est supposé être dans le creux du croissant (Du *et col.*, 2000) (Figure 1-13A). Deux types d'interfaces entrent en jeu pour l'oligomérisation de l'hexamère. Le premier, dit interface « AB » en référence au diagramme présenté en figure 1-13B, fait intervenir une paire de ponts salins et des interactions hydrophobes. Le deuxième, dit « BC », est essentiellement composé de liaison hydrogène. Les six monomères s'arrangent en anneau suivant les interfaces décrits ci-dessus pour former le complexe actif d'environ 120 kDa. Étonnamment, lors de la résolution de la structure du complexe Protease I, l'unité asymétrique ne contenait

que le trimère de Protease I. Il n'a donc pas été déterminé si l'assemblage est un dimère de trimère ou un trimère de dimère. Il a été récemment suggéré que Protease I pourrait être présente sous la forme d'un dodécamère de 240 kDa, cependant aucune donnée structurale n'est disponible (Zhan *et col.*, 2014).



**Figure 1-13 : Structure de la peptidase Protease I de *Pyrococcus horikoshii* (Code PDB : 1G2I).** A) Structure du monomère de PhPI représenté en ruban. H = hélice alpha ; S = feuillet bêta. Le site catalytique supposé est indiqué par un cercle rouge. B) Structure de l'hexamère de PhPI représenté en ruban. Les différents axes de symétrie sont représentés. C) Représentation en surface moléculaire de l'hexamère de PhPI coloré en fonction du potentiel électrostatique. Les sites actifs supposés sont indiqués par des flèches. (Tirée de Du *et col.* (2000))

Des homologues de Protease I existent dans les 3 domaines du vivant et leurs séquences sont bien conservées (Wagenfeld *et col.*, 1998; Takahashi *et col.*, 2001). Cependant le rôle de ces peptidases reste à ce jour indéterminé.

### 1.3.8 Pab87

La peptidase Pab87 est une exopeptidase à sérine appartenant à la famille S12 du clan SE. Elle est stéréospécifique et est active contre les acides aminés D. Elle fait partie de la famille de peptidase cubicO, dont les membres ont la particularité d'adopter une structure quaternaire en forme de cube. Elle a été caractérisée chez l'Archée hyperthermophile *Pyrococcus abyssi* (Delfosse *et col.*, 2009).

Le monomère de Pab87 de 50.4 kDa est constitué de deux domaines structuraux, un domaine N-terminal PRP ainsi qu'un domaine C-terminal de type lipocaline jouant un rôle essentiel dans l'octamérisation et dans la compartimentation du site actif de la peptidase (Figure 1-14A) (Delfosse *et col.*, 2009).

L'octamère de 400 kDa adopte la forme d'un cube dont les arêtes mesurent environ 100 Å, un pore central est présent le long de l'axe de symétrie 4 et plusieurs pores plus petits sont répartis le long de l'axe équatorial de la particule. L'octamère est formé de quatre dimères disposés en tête à queue organisé autour de l'axe de symétrie 4 de la peptidase et forme les arêtes du cube (Figure 1-14B et C). Chaque monomère interagit également, de façon secondaire, avec quatre autres sous-unités de l'octamère. Les interactions entre sous-unités sont de type électrostatique et de type liaison hydrogène entre la partie basse du domaine lipocaline d'un premier monomère et la région formée d'hélices  $\alpha$  du domaine PRP d'un second monomère (Delfosse *et col.*, 2009).

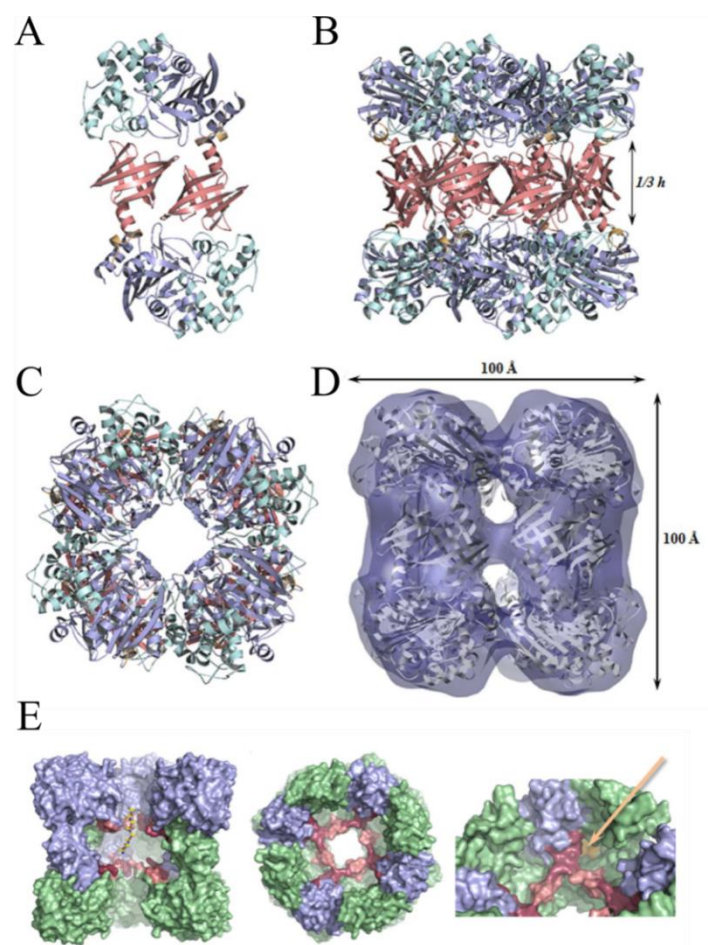


Figure 1-14 : Structure de la peptidase Pab87 (Code PDB : 2QMI). A) Dimère de Pab87 adoptant la conformation tête à queue représenté en ruban. Les domaines catalytiques et d'hélices  $\alpha$  sont représentés en bleu clair et cyan, l'hélice intermédiaire en orange et le domaine lipocaline en saumon. B, C) Octamère de Pab87 vu de côté (B) et de face (C) à l'axe perpendiculaire à l'axe 4 de symétrie de la particule. D) Fitting de la structure cristallographique dans l'enveloppe obtenue par microscopie électronique. E) Gauche : Vue interne de l'octamère de Pab87, le dimère du premier plan a été retiré. L'anneau tétramérique est coloré en bleu et vert. La couronne restreignant l'accès aux sites actifs est représentée en saumon. Milieu : Vue du dessus de l'octamère représentée en surface moléculaire. Pour une meilleure visualisation, les domaines N-terminaux ont été retirés. Droite : Vue identique à la précédente mais légèrement penchée afin de voir l'entrée du site actif représentée en orange. (Tirée de Delfosse 2009)

Un canal de forme carrée de 20 Å de côté traverse la particule le long de l'axe 4 de symétrie et rejoint une cavité centrale de 60 Å de diamètre formé par huit domaines C-terminaux. Chaque site actif fait face à la cavité interne d'environ 80 000 Å<sup>3</sup>. L'accès à la cavité catalytique est restreint au niveau des canaux d'entrée par une couronne formée de deux boucles issues de monomères opposés (Figure 1-14C et E) (Delfosse *et col.*, 2009).

Des homologues de Pab87 sont présents chez les Bactéries et les Archées. Tous les organismes possédant cette peptidase vivent dans des zones proches d'événements hydrothermaux. Il a été proposé que la peptidase Pab87 joue un rôle dans l'hydrolyse de débris cellulaire bactériens contenant des acides aminés D pour le métabolisme des cellules vivant dans ces zones extrêmes (Delfosse *et col.*, 2009).

### 1.4 SYNTHÈSE

---

Les peptidases cytosoliques ATP-indépendantes ont des rôles essentiels dans de nombreuses fonctions de la cellule. Ont été présentées ici celles capables de former des grands assemblages oligomériques auto-compartmentés. Le tableau 1-2 résume les principales caractéristiques de ces peptidases. Elles sont réparties dans les familles de peptidase S, C et M et possèdent des activités peptidases diverses, on peut cependant noter la présence d'une majorité d'exopeptidases. Ces complexes ont des types d'assemblage très divers. L'ensemble des peptidases présentées dont les structures cristallographiques sont disponibles sont représentées sur la Figure 1-15. Dans tous les cas présentés, sauf pour Tricorn et Protease I, l'oligomérisation permet la formation d'une ou plusieurs cavités à l'intérieur de la particule auxquelles est relié un réseau de canaux. C'est l'architecture de ce réseau interne qui permet un criblage efficace des substrats, protégeant ainsi le reste du protéome de l'activité hydrolytique des peptidases. D'autre part, la formation de ce réseau et des cavités internes associées regroupant plusieurs sites actifs augmente l'efficacité de la peptidase, notamment dans le cas de peptidase non-processive comme c'est le cas pour DegP. Ainsi, il ressort de nos analyses qu'une sélectivité précise et une efficacité augmentée sont deux propriétés qui présentent un grand intérêt dans le cas de peptidase impliquée dans des rôles physiologiques précis impliquant des peptides régulateurs. En effet, la peptidase est



alors capable de cibler précisément son activité sur le substrat d'intérêt qu'elle sera capable de cliver de façon efficace.

La plupart des complexes de peptidases semblent être bien conservés chez les procaryotes. DegP et Protease I le sont également chez les eucaryotes, on notera qu'il existe une forme proche de « trilobed protease » chez les eucaryotes également. Il serait intéressant de réaliser des études de phylogénie plus fines pour chacun de ces complexes afin de déterminer s'ils sont associés à des organismes ou à des environnements en particulier.

L'oligomérisation des peptidases permet de protéger le protéome de la cellule tout en augmentant l'efficacité de l'enzyme dans sa fonction. Cependant, la régulation de l'oligomérisation ou de l'activité des complexes peptidases auto-compartimentés dans la cellule est très peu étudiée. En effet, peu d'informations existent pour savoir si ces peptidases sont toujours présentes sous formes oligomériques dans le cytoplasme. Il en va de même en ce qui concerne leurs capacité à s'hétéro-oligomériser dans les cas où plusieurs versions de la protéine existent dans un même type cellulaire. La seule de ces peptidases pour laquelle un cas d'hétéro-oligomérisation existe est la forme de DegP présente chez les plantes. En effet, les peptidases Deg5 et Deg8 ont été observées sous la forme d'un complexe hétéro-oligomérique qui serait impliqué dans la réparation du photosystème II lorsqu'il est endommagé.

Le rôle précis des différentes peptidases "géantes" énergie indépendante dans la cellule reste dans la plupart des cas assez obscur. Ceci est étonnant mais s'explique par le manque d'études *in vivo*, en particulier dans des modèles cellulaires procaryotes.

L'étude des grands assemblages moléculaires en biologie structurale représente aujourd'hui un défi majeur. Les grands complexes qui ont été décrits dans ce chapitre ne représentent que ceux dont la structure a été caractérisée. On peut penser qu'il existe d'autres grands assemblages protéiques du même type qui n'ont pas encore été caractérisés. Dans le laboratoire ELMA, plusieurs grands assemblages qui pourraient être impliqués dans la protéolyse ont récemment été mis en évidence et sont en cours de caractérisation.

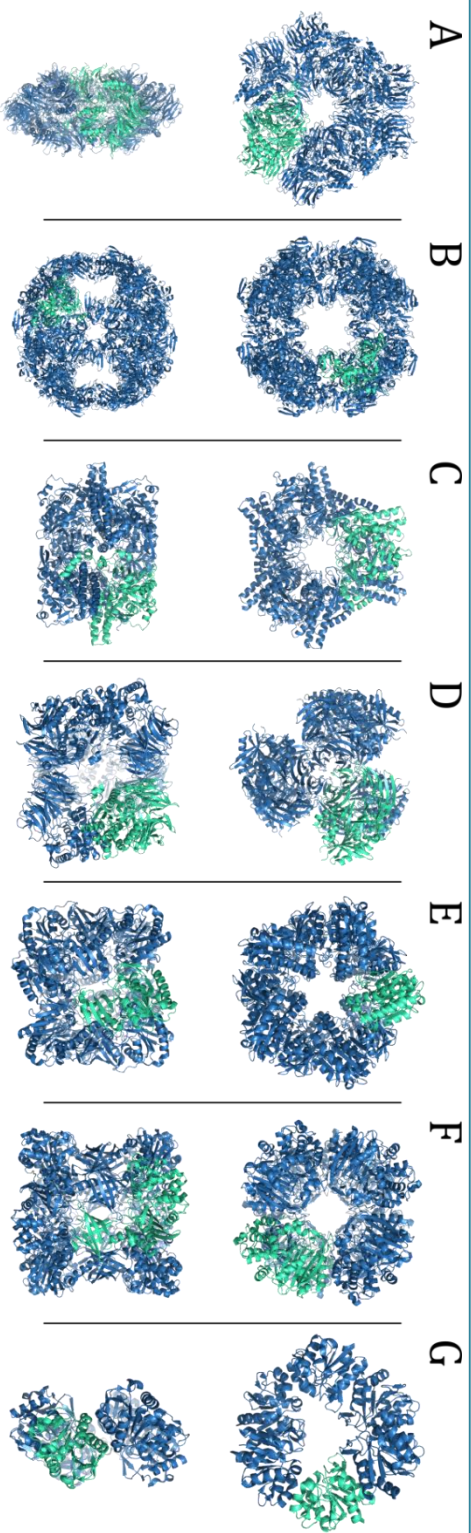


Figure 1-15 : Structures cristallographiques des peptidases formant de grands assemblages oligomériques représentés en ruban. Pour chaque peptidase, une première vue est proposée selon un axe de symétrie présente dans la suite (haut) et selon la perpendiculaire à cet axe (bas). A) Tricorn vue selon l'axe de symétrie 3 (Code PDB : 3CS0). B) DegP vue selon l'axe de symétrie 4 (Code PDB : 4HXE). C) Trilobed peptidase vue selon l'axe de symétrie 3 (Code PDB : 1GCB). D) Trilobed peptidase vue selon l'axe de symétrie 3 (Code PDB : 1GCB). E) DppA vue selon l'axe de symétrie 5 (Code PDB : 1H19). F) Pab87 vue selon l'axe de symétrie 4 (Code PDB : 2QMD). G) Protease I vue selon l'axe de symétrie 3 (Code PDB : 1G21).

Peptidase	famille MEROPS	Activité	Nombre de sous-unités dans l'assemblage	Forme	Activation	Rôle	Domaine	Hétéro-oligomérisation
Tricorn	S41	Endopeptidase de type chymotrypsine	6	Tricorne	Inconnu	Inconnu	Archée	Peut faire partie d'un complexe protéolytique
Tricorn	-	A priori identique	60 (20 hexamères)	Capsule	Inconnu	Inconnu	Archée	Inconnu
TPP II	S8	Tripeptidyl aminopeptidase (endopeptidase faible)	40 (20 dimères)	Double hélice	Via l'oligomérisation	Inconnu	Eucaryote	Inconnu
DegP	S1C	Endopeptidase + chaperonne	12 ou 24	Cage	Température / Allostérie	Heat shock protein	Eucaryotes, Bactérie, Archée	Oui, uniquement chez les plantes
Bleomycine	C1	Aminopeptidase	6	Anneau	Autocatalyse	Inconnu	Bactérie, Archée	Inconnu
Trilobed	S9	Deblocking acylaminoacyl peptidase + endopeptidase	6	Trèfle creux	Inconnu	Inconnu	Bactérie, Archée	Inconnu
DppA	M55	Aminopeptidase D-spécifique	10	Tonneau	Inconnu	Inconnu	Bactérie, Archée	Inconnu
Protease I	G56	Aminopeptidase (+ endopeptidase faible)	6	Anneau	Inconnu	Inconnu	Eucaryotes, Bactérie, Archée	Inconnu
Pab87	S12	exopeptidase D-spécifique	8	Cube	Inconnu	Inconnu	Bactérie, Archée	Inconnu

Tableau 1-2 : Résumé de différentes caractéristiques des peptidases présentées dans cette partie.

### 1.5 LES AMINOPEPTIDASES TET DE PYROCOCCLUS HORIKOSHII

---

Ce chapitre est consacré à la famille des peptidases TET, dernier type de peptidase géante auto-compartmentée ATP-indépendante que je vais présenter ici. Ce groupe de peptidases représente l'objet principal du travail réalisé durant ma thèse.

#### 1.5.1 Historique de la découverte des aminopeptidases TET.

Les tests d'activités d'une endoglucanase putative, exprimée chez l'Archée hyperthermophile *Pyrococcus horikoshii*, ont révélé une activité de clivage N-terminal sur une grande variété de substrats (Ando *et col.*, 1999). Cette enzyme présentait une forte similarité de séquence avec une protéine appelée « deblocking aminopeptidase », ou DAP, identifiée chez l'Archée *Pyrococcus furiosus* (Tsunasawa, 1998). Il a ensuite été montré que l'endoglucanase putative identifiée chez *P. horikoshii* présentait elle aussi une activité acylaminoacyl peptidase (Onoe *et col.*, 2002). Finalement, cette peptidase présentait des similarités structurales avec une grande aminopeptidase auto-compartmentée de près de 500 kDa possédant une large gamme de spécificité de substrat identifiée chez l'Archée halophile *Haloarcula marismortui* (Franzetti *et col.*, 2002a). Cette protéine a été découverte dans l'équipe de B. Franzetti dans le cadre de la recherche des complexes chaperons et protéases associés au stress salin chez les Archées halophiles extrêmes. Ce complexe présentait la particularité d'être composé de 12 sous-unités s'assemblant sous la forme d'un tétraèdre creux (Franzetti *et col.*, 2002b). Ce complexe très singulier a été nommé TET (pour tetrahedral peptidase) en référence au complexe peptidase TRI (voir section 1.3.1), qui elle, tire son nom de la forme du chapeau tricorne (Tamura *et col.*, 1996).

Cette découverte a déclenché un vif intérêt dans les laboratoires spécialisés dans la biologie structurale des peptidases auto-compartmentées, en particulier l'équipe de M. Groll au Max-Planck Institute à Munich et celle d'U. Bauman à Bern. La protéine TET halophile s'étant révélée impossible à produire sous forme recombinante, l'équipe de B. Franzetti à Grenoble s'est rapprochée du groupe d'A. Roussel à Marseille qui était engagé dans un programme de génomique structurale des cellulases de *P. horikoshii*. Les travaux préliminaires ayant déjà été réalisés (Durá *et col.*, 2005) sur PhTET1 (Porciero *et col.*, 2005) c'est sur cette protéine, pourtant la plus difficile du fait de sa capacité à



former également des 24-mers, que le groupe s'est concentré. Trois articles ont été ainsi publiés en peu de temps : deux sur la structure de PhTET2 par les équipes concurrentes et l'article sur PhTET1 qui comprend également une caractérisation du complexe octaédrique (Russo *et col.*, 2004; Borissenko *et col.*, 2005; Schoehn *et col.*, 2006). Depuis, en plus de la TET de *Haloarcula marismortui* (HmTET), les structures et les activités enzymatiques respectives des trois peptidases homologues de *P.horikoshii* ont été étudiées de manière approfondie à l'IBS

(Durá *et col.*, 2005; Schoehn *et col.*, 2006; Durá *et col.*, 2009; Rosenbaum *et col.*, 2011; Talon *et col.*, 2011; Franzetti *et col.*, 2013). Toutes sont effectivement des complexes qui présentent la structure quaternaire type de TET. Elles ont ainsi été nommées PhTET1, PhTET2 et PhTET3. L'ensemble de ce travail est résumé ci-dessous.

### 1.5.2 Fonction

Des metallopeptidases appartenant à une même famille enzymatique peuvent manipuler des substrats différents, voire présenter des fonctions différentes. Pour caractériser le système TET de *P. horikoshii*, l'activité des trois peptidases TET a été étudiée en utilisant toute une variété de substrats chromogéniques ou fluorogéniques, de tailles et de compositions différentes. Ces substrats sont disponibles dans le commerce (Bachem, Sigma, ...). Des profils de digestion de peptides synthétiques de tailles et de compositions diverses ont également été analysés par HPLC en phase inverse (voir section 2.1.1.2 et chapitre 4). Ces études montrent clairement que les trois peptidases TET sont des aminopeptidases. C'est à dire des enzymes qui éliminent un par un les acides aminés à partir des extrémités N-terminales des polypeptides. A ce jour aucune activité di- ou tripeptidyl peptidase, N-deblocking, endopeptidase ou carboxypeptidase n'a pu être détectée. L'analyse des cinétiques de libération d'acides aminés durant ces expériences a également permis de montrer que les peptidases TET ne sont pas processives (Durá 2005), c'est à dire que le substrat est relâché par l'enzyme après chaque coupure.

### 1.5.3 Mécanisme enzymatique

Les peptidases TET de *P.horikoshii* sont toutes rangées dans la famille M42 du clan MH selon la classification MEROPS. Les résidus impliqués dans la catalyse ainsi que l'étude des sites actifs dans les structures cristallographiques des 3 TET de *P.horikoshii*

montrent une grande conservation. Ceci permet de décrire le mécanisme catalytique de l'enzyme.

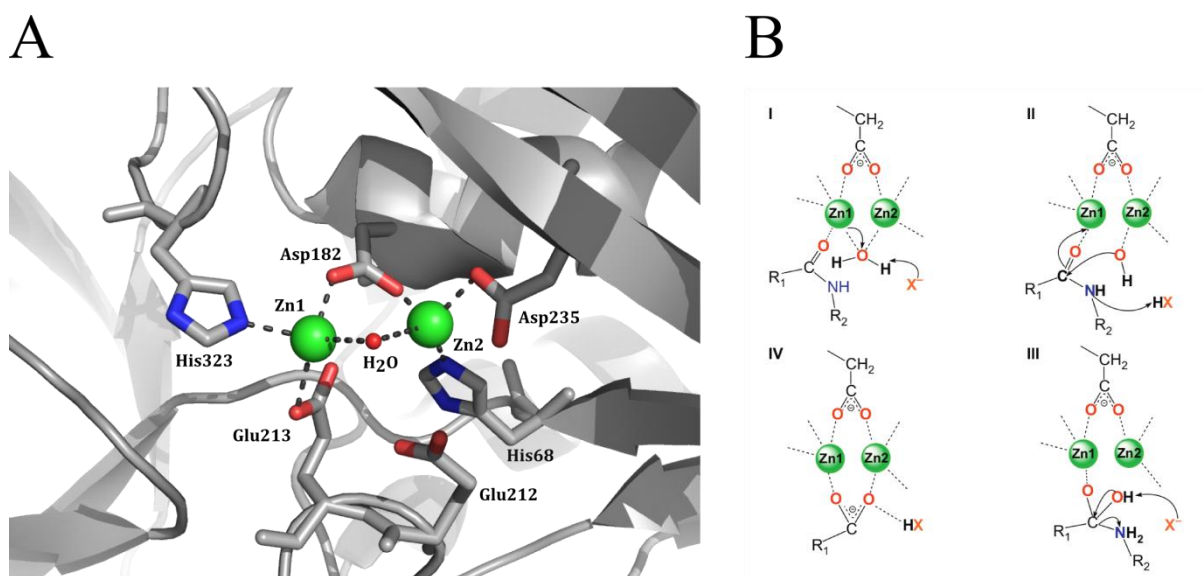


Figure 1-16 : Mécanisme catalytique des aminopeptidases TET. A) Représentation du site actif de l'aminopeptidase PhTET2 (Code PDB: 1Y0R). Les résidus impliqués dans la coordination des ions catalytiques et dans la catalyse sont représentés en bâtonnets. Les ions zinc sont représentés par des sphères vertes, et l'atome d'oxygène de la molécule d'eau est représenté par une sphère rouge. Dans le cas de PhTET2, l'ion Zn1 est coordonné par les résidus suivants : Asp182, Glu 213 et His 323. L'ion Zn2 est aussi coordonné par l'Asp182 ainsi que par les résidus His68 et Asp235. B) Mécanisme catalytique des aminopeptidases à site bimétalliques (Tirée de Auld (2004)). Le 'X' est un catalyseur acide/base du site actif, dans le cas de TET, il s'agit du résidu Glu212.

Le domaine catalytique des aminopeptidase TET est très similaire à celui de la peptidase AAP (pour *Aeromonas proteolyticus* aminopeptidase), l'enzyme type du clan MH (Russo *et col.*, 2004; Schoehn *et col.*, 2006). Plusieurs études ont été réalisées sur le mécanisme catalytique de cette peptidase (Bennett *et col.*, 1997; Chen *et col.*, 1997; De Paola *et col.*, 1999; Stamper *et col.*, 2001; Bzymek *et col.*, 2004). Cette peptidase, ainsi que toutes celles du clan MH, possède deux cations divalents dans le site actif (Chevrier *et col.*, 1994). Dans le cas des peptidases TET, des ions cobalt ou zinc ont été modélisés dans les sites actifs (Russo *et col.*, 2004; Schoehn *et col.*, 2006).

Dans le cas de l'AAP, comme dans le cas de TET, ce sont le groupe carboxylate d'un résidu acide aspartique ainsi qu'une molécule d'eau qui lient les deux ions catalytiques (Chevrier *et col.*, 1994; Russo *et col.*, 2004). Chaque ion est également coordonné par l'azote Nε2 d'un résidu histidine. De plus, l'ion 1 est coordonné par l'oxygène Oε1 d'un résidu acide glutamique alors que l'ion 2 est coordonné par l'oxygène Oδ1 d'un résidu

acide aspartique. Un résidu acide glutamique est présent dans le site actif, il est impliqué dans la catalyse pendant laquelle il forme une liaison hydrogène avec la molécule d'eau.

Le site actif de PhTET2 est représenté en figure 1-16A, le mécanisme général d'hydrolyse des aminopeptidases à site bimétalliques telles que TET est présenté sur la figure 1-16B. Lors de la première étape de ce mécanisme, un groupe carbonyle du résidu  $P_1$  du peptide interagit avec l'ion Zn1. Le résidu Glu212 agit alors comme une base en captant un proton de la molécule d'eau fixée à l'ion Zn2. L'hydroxyde ainsi formé peut attaquer le groupe carbonyle du peptide scissile pour former l'intermédiaire tétraédrique. La forme protonnée du résidu Glu212 peut alors agir comme un acide catalytique en cédant un proton à l'azote du groupe amide du résidu  $P'_1$  du peptide provoquant la rupture de la liaison peptidique (Auld, 2004).

### 1.5.4 Substrats

L'unique peptidase TET de *H. marismortui* (HmTET) est capable d'hydrolyser depuis l'extrémité N-terminale des peptides dont la taille atteint 40 résidus (Franzetti *et col.*, 2002b). Les études réalisées sur les 3 TET de *P. horikoshii* suggèrent que les particules TET sont dans l'incapacité de dégrader des chaînes polypeptidiques repliées.

HmTET est capable de cliver une large gamme d'acides aminés. Par contre, chacune des trois TET de *P. horikoshii* possède une spécificité de substrats différente.

PhTET1 présente la particularité de ne pas pouvoir hydrolyser les substrats chromogéniques classiques tel que X-pNA ou X-AMC, l'utilisation d'un protocole adapté a donc été nécessaire pour mesurer l'hydrolyse de l'acide aminé N-ter. Une recherche des substrats spécifiques à partir de mélanges de peptides a été réalisée au laboratoire. Pour cela un hydrolysats de caséine a été incubé avec l'enzyme. Après séparation du mélange peptidique sur colonne HPLC, les peptides substrats ont été identifiés et purifiés et leur séquence N-ter a été déterminée par dégradation Edman. Ceci a permis de montrer que PhTET1 possède une spécificité de substrat étroite dirigée essentiellement contre les acides aspartique et glutamique. Le peptide synthétique EGI a ainsi été retenu comme substrat type de PhTET1 (Franzetti *et col.*, 2013). Ainsi PhTET1 est annotée comme une aspartyl/glutamyl-aminopeptidase.

PhTET2 est une Leucyl-aminopeptidase montrant une préférence pour les résidus non chargés, elle possède la plus large spécificité chez les PhTET puisqu'elle est capable d'hydrolyser des peptides dont le N-ter est constitué soit d'une leucine, une méthionine, une isoleucine, une sérine ou une alanine. L'activité maximale enregistrée lors des tests d'activité l'a été contre le substrat chromogénique Leu-pNA (Durá *et col.*, 2005).

Enfin PhTET3 est une lysyl-aminopeptidase hydrolysant les résidus chargés positivement. Le maximum d'activité de cette aminopeptidase a été enregistré lors de l'utilisation de substrat Lys-pNA ou Arg-pNA, cependant elle est aussi capable d'hydrolyser les résidus Glu ou Leu, dans une moindre mesure (Durá *et col.*, 2009).

Les aminopeptidases TET de *P.horikoshii* présentent donc des affinités de substrats complémentaires, non redondantes. Pour cette raison il a été proposé que ces trois enzymes forment un système intégré de destruction des peptides chez les Archées hyperthermophiles. Le travail réalisé dans le cadre de cette thèse nous a amené à revoir cette hypothèse.

### 1.5.5 Co-facteurs et paramètres enzymatiques optimaux

L'effort d'identification des substrats spécifiques à chacune des 3 TET de *P. horikoshii* a permis de déterminer les conditions optimales pour réaliser leur fonction. Les trois peptidases ont une activité optimale à des températures comprises entre 95 et 100 °C. En dessous de 70 °C, elles ne montrent plus qu'une activité comprise entre 20 et 40 % de leur activité optimale. Toutes trois montrent une grande thermostabilité puisqu'elles ont des demi-vies d'environ 3 à 4h à 90 °C.

En utilisant le substrat Glu-Gly-Ile à 90 °C, l'aminopeptidase PhTET1 montre une activité optimale à pH = 7.4 dans un tampon CHES, plus de 50 % de l'activité optimale est mesurée entre les pH = 7 et 7.8. Dans le cas de PhTET2, l'activité optimale de l'aminopeptidase a été mesurée à pH = 7.5 dans un tampon PIPES à 80 °C contre le substrat Leu-pNA, elle montre encore plus de 60 % de son activité à pH 6 ou 9. L'aminopeptidase PhTET3 montre une activité optimale à pH = 7.4 dans un tampon TAPS à 85 °C contre le substrat Lys-pNA, 50 % de son activité optimale est mesurable à pH 6.5 et 8.

Les aminopeptidases PhTET1, PhTET2 et PhTET3 sont toutes fortement activées en présence de cobalt à environ 1 mM. PhTET1 est la plus activée par la présence d'ions cobalt. Elles sont également toutes les trois inhibées en présence d'ion  $Zn^{2+}$  ou  $Mn^{2+}$  à 0,5 mM. Étant des métallopeptidases, elles sont fortement inhibées par la présence d'EDTA dans le tampon.

Enfin, une étude réalisée par Rosenbaum *et col.* (2011) sur l'aminopeptidase PhTET3 montre le rôle déterminant du pH et des ions métalliques pour la stabilité de l'édifice dodécamérique. En effet, au-delà de pH = 8, la désoligomérisation du complexe est observée. À pH = 11 seul 50 % de la protéine est encore sous forme dodécamérique après une incubation d'une nuit. Une incubation pendant huit jours en présence d'EDTA à 20mM conduit à la désoligomérisation quasi-totale du complexe. Il est intéressant de noter que ce dernier phénomène est en partie réversible (Rosenbaum *et col.*, 2011).

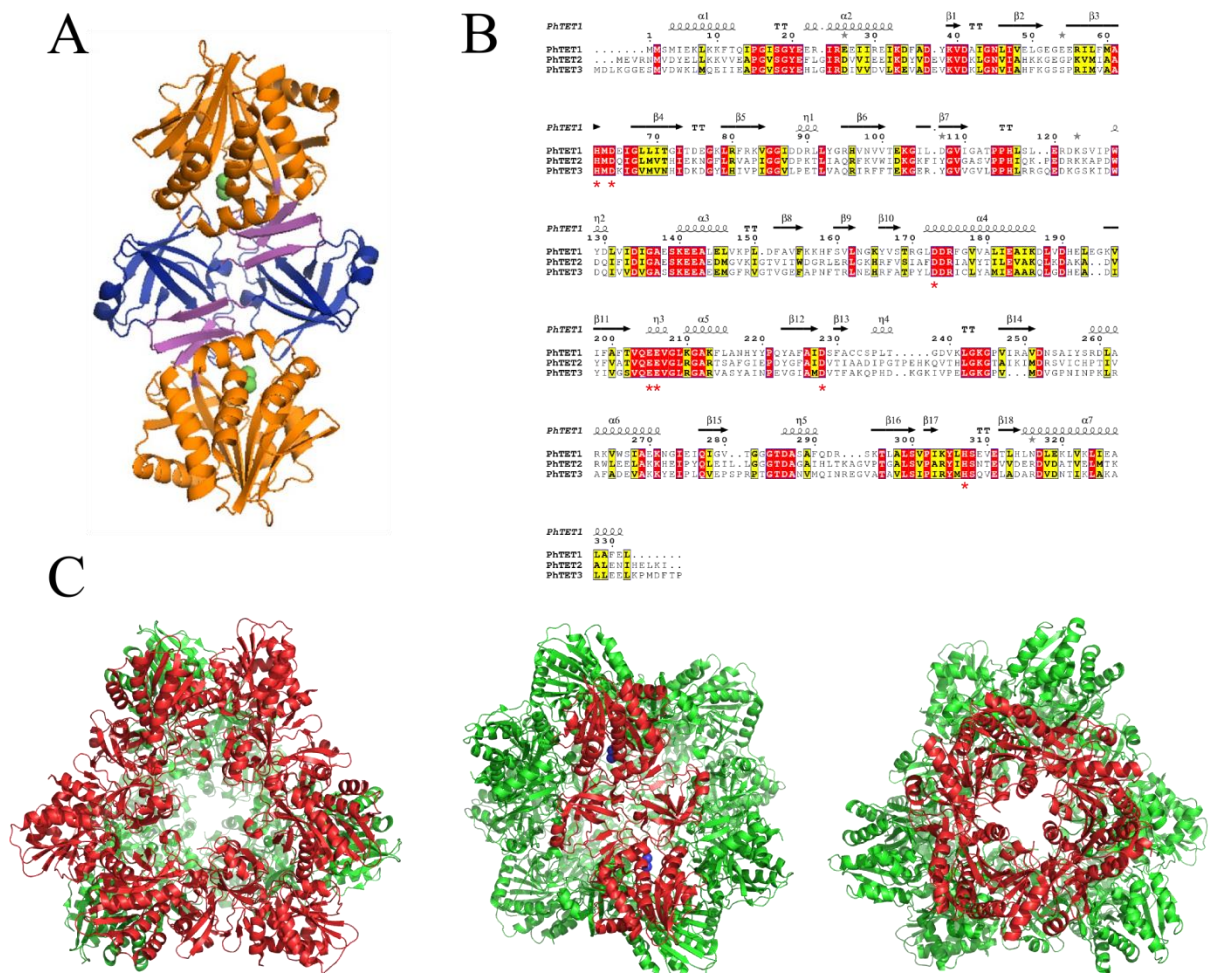
### 1.5.6 Structures

Les structures cristallographiques des 3 TET de *P.horikoshii* ont été résolues et caractérisées. Dans chacune des structures disponibles, on trouve le monomère de TET dans l'unité asymétrique et il est possible d'obtenir la structure du dodécamère à partir de la symétrie cristalline.

Le monomère de PhTET1 est composé de 332 résidus et a un poids moléculaire d'environ 37 kDa. Il est constitué de 2 domaines : un large domaine catalytique comprenant les résidus 2 à 65 et 156 à 331, ce domaine porte le site bimétallique; et un domaine de dimérisation plus petit compris entre les résidus 66 et 155 (Figure 1-17). Les monomères de PhTET2 et PhTET3 adoptent une conformation très proche, mais pas identique. En effet les monomères de PhTET2 et PhTET3 sont plus longs, respectivement de 21 et 22 résidus, leurs poids moléculaires sont d'environ 39kDa (Durá *et col.*, 2009), la structure générale de la protéine restant la même.

Un dimère de TET est formé par l'association du domaine de dimérisation d'un monomère (en bleu sur la figure 1-17A) avec une partie du domaine catalytique d'un deuxième monomère (en mauve sur la figure 1-17A) (Schoehn *et col.*, 2006). Il a été proposé que le dimère soit la structure de base pour l'assemblage du dodécamère.

Six dimères s'assemblent pour former la particule dodécamérique. Le complexe à 12 sous-unités a un poids moléculaire d'environ 400 kDa et mesure près de 135 Å de diamètre, il présente la forme générale d'un tétraèdre.



**Figure 1-17 : Structure cristallographique de PhTET1.** A) Représentation de la structure du dimère de PhTET1 en ruban. (tirée de Schoehn *et col.* (2006)) B) Alignement de séquence multiple des aminopeptidases PhTET1, PhTET2 et PhTET3. Les structures secondaires indiquées correspondent à la structure de PhTET1. Les résidus du site actif sont indiqués par une étoile C) À gauche : Représentation de la structure du dodécamère de PhTET1 vue selon l'axe de symétrie d'ordre 3 du tétraèdre. Les dimères formant une des faces du tétraèdre sont colorés en rouge. Au centre : Représentation de la structure du dodécamère de PhTET1 vue selon l'axe de symétrie d'ordre 2. Un dimère, constituant l'arête du tétraèdre est visible, les ions métalliques des sites actifs de ce dimère sont représentés par des sphères bleues. À droite : Représentation de la structure du dodécamère de PhTET1 vue selon l'axe de symétrie d'ordre 3 mais du côté opposé au premier. Vu de l'apex du tétraèdre, le trimère de PhTET formant l'apex est coloré en rouge. (inspirée de Durá *et col.* (2009))

La structure du site actif est typique des M42, composé des résidus H61XD63, D172, E204E205, D227 et H306 qui coordonnent deux ions métalliques  $\text{Co}^{2+}$  et permettent l'activité d'hydrolyse des peptides.



Chaque apex du tétraèdre est formé par un trimère de TET arrangé en cercle. Ces trois sous-unités forment un pore de 7 Å diamètre chez PhTET1 (Schoehn *et col.*, 2006) (Figure 1-17C). Ces pores, situés aux apices du complexe sont fermés chez PhTET2 et PhTET3 où ils sont bloqués par les résidus F224 et Y227 respectivement (Borissenko *et col.*, 2005; Durá *et col.*, 2009). Chaque arête du tétraèdre est constituée d'un dimère de TET (Figure 1-17C), trois dimères arrangés en triangle forment donc une face de l'édifice où se trouve au centre un pore plus grand (Figure 1-17C) (Borissenko *et col.*, 2005; Schoehn *et col.*, 2006; Durá *et col.*, 2009).

### 1.5.7 Mode d'action

L'analyse de la partie interne des particules TET révèle une compartimentation complexe de l'édifice. Les structures générales des complexes dodécamériques sont quasi identiques entre les 3 TET caractérisées.

Les grands pores situés sur chacune des quatre faces du tétraèdre donnent accès à quatre canaux qui s'entrecroisent au centre de la particule et constitueraient des pores d'entrées (Schoehn *et col.*, 2006). Ce pore d'entrée est visible sur la figure 1-17C à gauche. La distribution des charges électrostatiques dans le canal est très contrastée et permet le guidage du substrat (Figure 1-18A) (Durá *et col.*, 2009). Ces canaux sont reliés à des poches catalytiques formées par les trimères situés aux apices du tétraèdre, chacune des quatre poches catalytiques de TET compte donc 3 sites actifs. Un trimère formant une poche catalytique est visible sur la figure 1-17C à droite. Ces poches catalytiques s'ouvrent sur l'extérieur de la particule *via* les petits pores situés aux apices du complexe. L'ensemble des quatre canaux s'entrecroisent et conduisent aux poches catalytiques comme montré sur la Figure 1-18B. Ces petits pores ne sont pas assez larges pour permettre le passage d'un long peptide, il a été proposé que leur rôle soit plutôt l'expulsion des acides aminés libres après leur hydrolyse dans la poche catalytique (Schoehn *et col.*, 2006). Ces pores sont bloqués chez PhTET2 et PhTET3, il a été proposé que l'hydrolyse de l'extrémité N-ter induise un changement conformationnel qui conduit à l'ouverture du pore permettant ainsi la sortie de l'acide aminé libre (Durá *et col.*, 2009). Une autre hypothèse proposée pour PhTET2 par Borissenko *et col.* (2005) serait que les produits soient reconduits dans la cavité centrale et soient expulsés via de petits pores situés autour du grand pore d'entrée (Borissenko *et col.*, 2005).

L'examen des surfaces des chambres catalytiques et des poches qui accueillent le substrat permet aussi d'expliquer les raisons de la différence de spécificité entre les 3 TET. Ainsi la poche catalytique adjacente au site actif de PhTET3 est chargée négativement et favorise l'activité de celle-ci contre les résidus chargés positivement comme la Lysine.

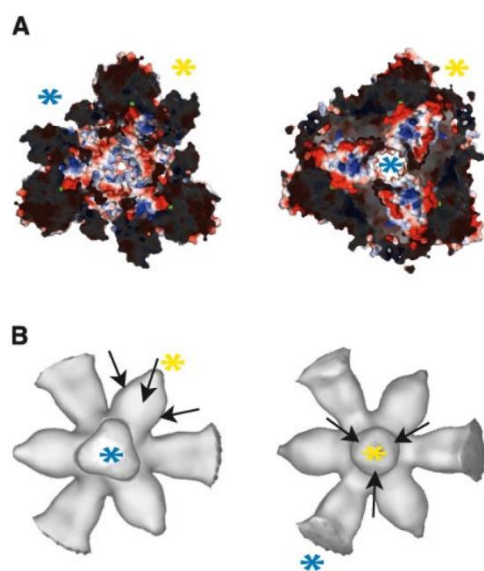


Figure 1-18 : Intérieur de la particule PhTET1. A) Deux coupes transversales de l'édifice tétraédrique représenté en surface. La surface de la particule est colorée suivant le potentiel électrostatique négatif (bleu) ou positif (rouge). A gauche : vu depuis un pore d'entrée menant à une poche catalytique chargée positivement. A droite: vu opposée. Les ions cobalt sont colorés en vert. Les étoiles bleues et jaunes indiquent respectivement où se trouve le large pore d'entrée (face du tétraèdre) et le petit pore de sortie (apex du tétraèdre). B) Représentation des cavités internes du complexe dodécamérique obtenue à partir de reconstruction du complexe réalisé par microscopie électronique (Schoehn *et col.*, 2006). Les flèches indiquent la position des sites actifs dans une poche catalytique. Les étoiles bleues et jaunes ont la même signification qu'en A. (Schoehn *et col.*, 2006)

La poche catalytique de PhTET2 est au contraire relativement neutre favorisant l'activité de la peptidase contre les petits résidus hydrophobes ou les résidus neutres comme la Leucine. Enfin, la poche catalytique de PhTET1 est chargée positivement, favorisant l'activité de l'aminopeptidase contre les résidus chargés négativement comme l'aspartate et le glutamate.

### 1.5.8 Rôle physiologique

Les aminopeptidases représentent un groupe d'enzymes qui peuvent être présentes dans le cytoplasme, associées à la membrane cellulaire ou bien sécrétées en dehors de la cellule (Gonzales *et col.*, 1996). Le domaine catalytique des TET présente une forte similarité de structure avec celui de l'aminopeptidase monomérique sécrétée AAP de la famille M28 du clan MH (voir section 3.1). Cependant la peptidase TET d'*Haloarcula marismortui* n'a été détectée que dans les fractions intracellulaires lors de sa découverte. Il a donc été proposé que le domaine de dimérisation des TET soit le résultat d'une insertion qui a eu lieu durant leur évolution et qui ait permis l'auto-compartimentation



de cette famille de peptidases autorisant ainsi sa localisation cytoplasmique (Franzetti *et col.*, 2013).

Le rôle physiologique spécifique des TET reste aujourd'hui inconnu. Considérée indépendamment les unes des autres, leurs spécificités de substrat plus ou moins fortes pourraient contribuer à modifier la composition des peptides dotés de propriétés physiologiques spécifiques. PhTET1 présentant une forte spécificité de substrat, pourrait être impliquée dans la dégradation ou la maturation de protéines (N-ter exposés) ou de peptides spécifiques dans la cellule. Il est aussi possible que PhTET1 soit mobilisée en condition de carence pour fournir des acides aminés comme le Glutamate qui peut être directement utilisé pour alimenter le cycle de Krebs. Considérées ensemble comme un système intégré de destruction des peptides, elles pourraient jouer un rôle dans la dégradation non spécifique de peptides dans le cytoplasme. Chez *Pyrococcus horikoshii*, elles pourraient alors contribuer au métabolisme en dégradant les polypeptides importés du milieu extracellulaire (Ward *et col.*, 2002). Cette dernière hypothèse est soutenue par une étude montrant qu'un gène codant pour un homologue de PhTET1 chez *Pyrococcus furiosus* est fortement activé lorsque les cellules sont cultivées dans un milieu enrichi en peptides (Schut *et col.*, 2003).

De plus, les peptidases TET pourraient jouer un rôle dans l'étape finale de la protéolyse intracellulaire en dégradant les peptides de 6 à 20 résidus produits par l'activité endopeptidase des différentes machines protéolytiques ATP-dépendantes comme le protéasome ou Lon ou Clp. Un des objectifs de cette thèse est d'apporter des informations pour mieux comprendre le rôle des aminopeptidases TET en étudiant les relations oligomérisations-fonctions et les interactions qui pourraient exister entre les différents membres de cette famille de peptidases.

### 1.6 LES ARCHEES: ORGANISMES MODELES POUR LES ETUDES DES PEPTIDASES TET

---

#### 1.6.1 Le troisième domaine du vivant

Avant-propos: Le travail de thèse porte sur une machinerie cellulaire issue d'Archées. Pour apprécier l'intérêt général en biologie des études sur les TET archéennes il est

nécessaire de présenter le domaine des Archées qui reste encore mal connu de la communauté scientifique.

### 1.6.1.1 Découverte des Archées

*Pyrococcus horikoshii*, l'organisme d'origine des protéases TETs étudiées dans ce projet, est une Archée. La découverte des Archées est issue des travaux de Carl Woese et George Fox dans les années 70 (Woese *et col.*, 1977). Leurs travaux, basés sur l'analyse des séquences de l'ARN ribosomal, ont permis de montrer qu'il existait au sein des procaryotes deux ensembles d'organismes bien distincts (Woese *et col.*, 1990). Le nouveau groupe ressemblait aux microorganismes méthanogènes, supposés posséder un métabolisme "primitif". Pour cette raison, ce nouveau groupe a été baptisé "Archaeobactéries". Cette découverte est une avancée décisive dans la compréhension du monde du vivant. En effet, les études génomiques liées à l'accroissement du nombre de génomes disponibles, conjuguées à des études phylogénomiques ont permis de montrer clairement que les "Archaeobactéries" représentent en fait un troisième domaine du vivant, les Archées, bien distincts des Bactéries et des Eucaryotes (Gribaldo *et col.*, 2006). Il apparaît de plus que LACA (Last Archaeal Common Ancestor) s'est distingué des deux autres domaines très tôt et est apparenté à l'ancêtre lointain des Eucaryotes (Figure 1-19).

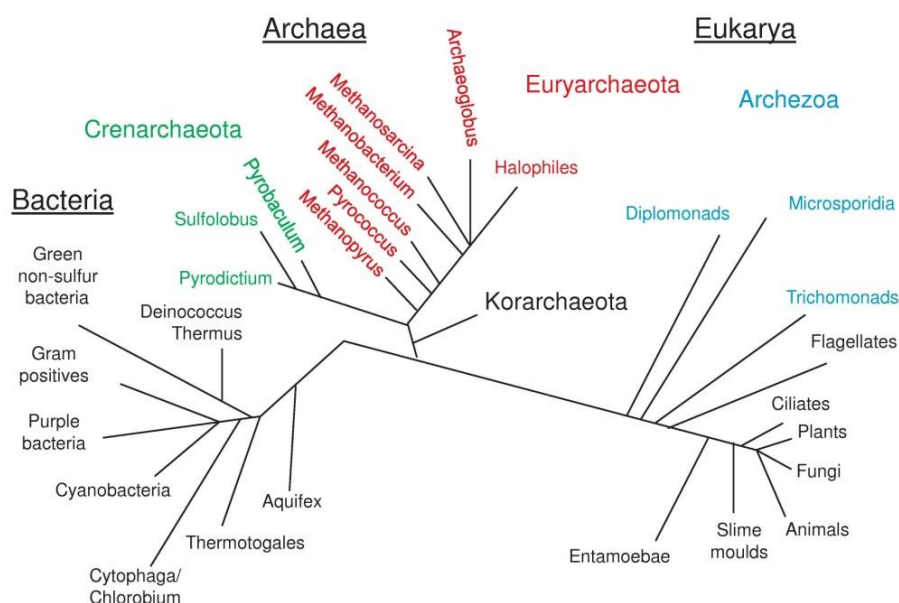


Figure 1-19 : L'arbre phylogénétique universel du vivant proposé par Woese *et col.* (1977).

### 1.6.1.2 Identité des Archées

L'analyse des génomes et du métabolisme des Archées révèle qu'elles possèdent des caractéristiques à la fois eucaryotes et bactériennes, avec toutefois certaines spécificités qui leur sont propres (Whitman *et col.*, 1999).

Ainsi, toutes les machines responsables des grands processus cellulaires: transcription, traduction, réplication et réparation de l'ADN (primases, hélicases, etc...), métabolisme de l'ARN, chaperonnes, protéasomes présentent de fortes similitudes avec les machines eucaryotes (Werner *et col.*, 2011). Leurs complexités sont cependant réduites. Par exemple, le protéasome 20S des Archées ne possède qu'entre deux et quatre sous unités différentes, contre 14 pour son homologue eucaryote (Maupin-Furlow *et col.*, 2006). Par contre, la présence d'unités transcriptionnelles polycistroniques ou encore l'organisation de leur chromosome rapproche les Archées des Bactéries (Samson *et col.*, 2011).

Dans le domaine du métabolisme, les annotations des génomes basés sur les similarités de séquences, ainsi que les études portant sur des voies métaboliques ou sur des enzymes connues indiquent que les Archées utilisent des systèmes qui sont plutôt de type bactérien. Cependant, il faut remarquer que beaucoup de ces enzymes ont pu être obtenues par transfert de gène horizontal.

Par ailleurs, les Archées possèdent des caractéristiques qui leur sont propres (Capes *et col.*, 2012). Ainsi toutes les Archées possèdent une membrane phospholipidique constituée de chaînes isoprénoïdes reliées par une liaison éther à un glycérol-1-phosphate alors que chez les eucaryotes et les bactéries, il s'agit de liaison ester et de glycérol-3-phosphate (Cavicchioli, 2011). De plus les Archées ne possèdent pas de paroi de peptidoglycane (Oger *et col.*, 2013) mais des parois de glycoprotéines. Ces organismes sont les seuls capables de produire du méthane à partir de CO<sub>2</sub> et d'H<sub>2</sub> (Welte *et col.*, 2011). L'accumulation récente de données de métagénomique environnementale ainsi que des travaux sur les métabolismes des communautés microbiennes marines ou terrestres ou encore l'analyse des voies métaboliques suggèrent l'existence d'enzymes inconnues, propres aux Archées, qui restent à décrire (Ferrer *et col.*, 2007; Sato *et col.*, 2011; Hofer, 2013). Enfin, la virologie des Archées

révèle les Archées et les Bactéries sont infectés par des virus appartenant aux mêmes familles et possèdent des plasmides semblables (Forterre, 2013).

### 1.6.1.3 Importance des Archées dans les écosystèmes

Les Archées ont longtemps été associées à l'extrémophilie. Cependant, les travaux récents de microbiologie environnementale révèlent que les Archées sont présentes dans tous les écosystèmes (Reeve *et col.*, 2011). Il est néanmoins vrai que les Archées dominent souvent les populations microbiennes qui se développent dans des conditions physico-chimiques considérées, de notre point de vue, comme "extrêmes" ou "a priori" hostiles à la vie. Il s'agit des très hautes températures que l'on trouve dans les sources hydrothermales, souvent associées à des pH acides ou à de fortes pressions hydrostatiques encore dans des environnements dans lesquels la salinité dépasse 4M, comme dans les grands lacs salés (Pikuta *et col.*, 2007). Dans le cas des hautes températures la prévalence des Archées est peut être due aux propriétés de leur membrane (Stetter, 1999). De plus, les études génomiques sur les communautés extremophiles des cheminées hydrothermales ou des lacs salés révèlent parfois une plus grande diversité d'Archées par rapport aux Bactéries (Takai *et col.*, 1999; Daffonchio *et col.*, 2006; Andrei *et col.*, 2012). L'accès aux génomes de microbes non cultivables, qui représentent plus de 90% des communautés microbiennes, montre également une biodiversité archéenne insoupçonnée (Reeve *et col.*, 2011; Haegeman *et col.*, 2013) Ce sont ces travaux qui depuis dix ans ont permis de montrer que les Archées sont présentes dans les sols, le permafrost, les sédiments, les océans et représentent même une partie du microbiome humain (Reeve *et col.*, 2011; Lepage *et col.*, 2013). Ces Archées "mésophiles" appartiennent principalement aux organismes méthanogènes du phylum *Euryarchaeota* ou au phylum des *Thaumarchaeota* (Chaban *et col.*, 2006). Etonnamment, bien qu'on trouve des Archées symbiote des eucaryotes (éponges marines, par exemple), aucune Archée pathogène n'a été découverte à ce jour. Elles sont cependant souvent indirectement associées à des pathologies dentaires ou à l'obésité (Conway de Macario *et col.*, 2009). Dans leur ensemble on estime que les Archées représenteraient 20 % de la biomasse terrestre (DeLong *et col.*, 2001). Elles doivent donc être considérées, non plus comme des organismes marginaux, mais comme des acteurs à part entière de l'évolution du vivant et des cycles géochimiques de la planète.

### 1.6.1.4 Diversité des Archées

Pour comprendre les travaux sur la diversité des peptidases TET présentés dans le chapitre 3 de la thèse, il est nécessaire d'introduire la phylogénie des Archées. L'arbre phylogénétique proposé par Brochier *et col.* (2005) est présenté en figure 1-20. Il existe deux principaux phylums chez les Archées : les *Crenarchaeota* et les *Euryarchaeota* (Gribaldo *et col.*, 2006; Brochier-Armanet *et col.*, 2011; Gribaldo *et col.*, 2011). Les *Crenarchaeota* regroupent des organismes thermophiles extrêmes aérobies ou anaérobies, souvent acidophiles également. Ils comprennent 3 ordres : les *Thermoproteales*, les *Desulfococcales* et les *Thermo-sulfolobales*. Les *Euryarchaeota* sont plus diversifiées et regroupent 9 ordres. Ceci reflète probablement une plus grande diversité de niches écologiques. On y trouve en effet des thermophiles mais aussi des mésophiles, halophiles ainsi que des aérobies ou anaérobies méthanogènes. *Pyrococcus horikoshii* est une *Euryarchaeota*. La microbiologie environnementale a récemment révélé l'existence d'au moins trois phylums supplémentaires: les *Thaumarchaeota* et les *Aigarchaeota* qui constituent une lignée ancienne qui regroupe des organismes présents de manière abondante dans les sols, l'eau douce et les océans (Lloyd *et col.*, 2013). Certains de leurs membres sont capables d'oxyder l'ammoniaque (Ammonia Oxidizers: AMO) et joueraient un rôle clé dans le cycle de l'azote (Jarrell *et col.*, 2011). Les *Korarchaeota* ont quant à elles été récemment découvertes dans des sources hydrothermales du parc de Yellowstone (Miller-Coleman *et col.*, 2012). Enfin, les *Nanoarchaeota* sont des cellules qui vivent en association stricte avec des *Desulfococcales*. Avec 600 gènes environ, elles possèdent un des plus petits génomes connus (Brochier *et col.*, 2005).

Les méthanogènes se trouvent dans des milieux diversifiés : des sédiments au tractus intestinal et jusqu'à des milieux plus extrêmes: permafrosts arctique et antarctique et sources d'eau chaude. Ils utilisent les produits du métabolisme de bactéries, sulfato-réductrices: CO<sub>2</sub>, Hydrogène, acétate, formate, méthanol et produisent du méthane (Welte *et col.*, 2011).

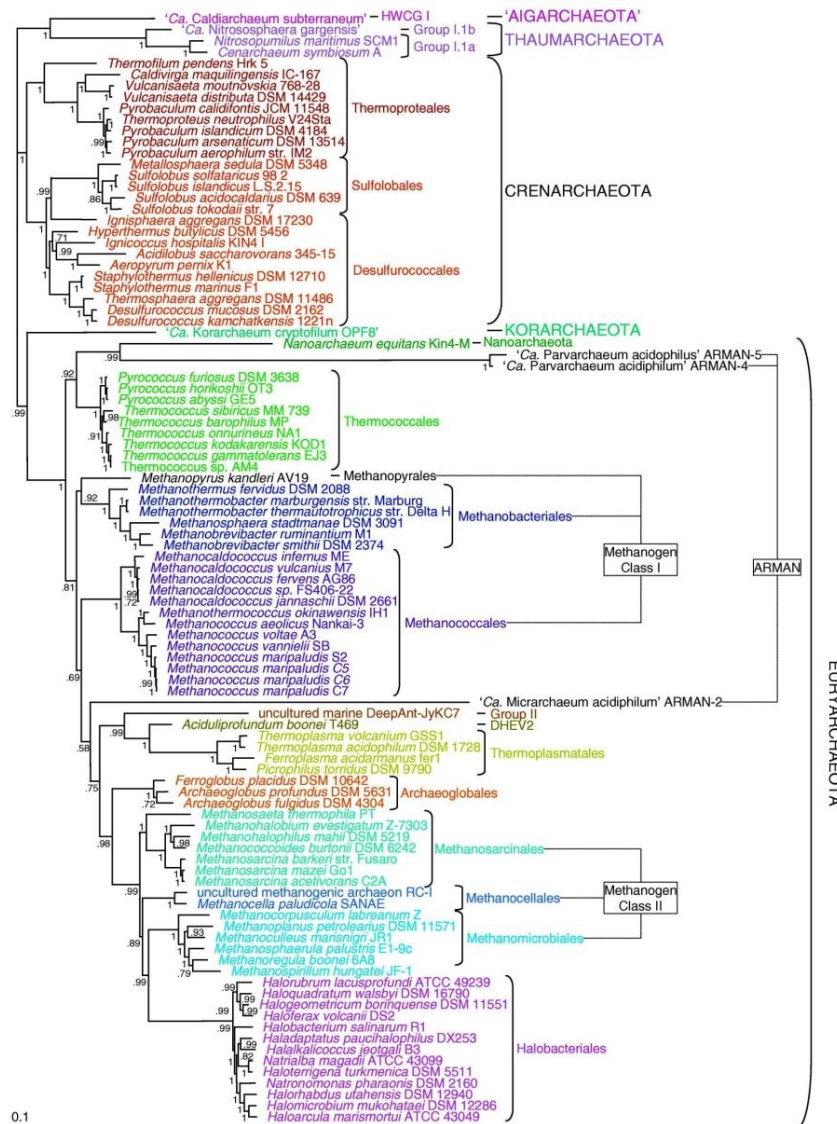


Figure 1-20 : Arbre phylogénétique et classification des Archées proposée par Brochier-Armanet et col. (2011).

## 1.6.1.5 Physiologie des Archées

Pour discuter du rôle des peptidases TET vis à vis du métabolisme ou d'une éventuelle adaptation environnementale il est nécessaire de présenter les 3 différents types de métabolismes utilisés par les Archées. Les Archées se caractérisent par une très grande plasticité métabolique entre les espèces mais aussi pour un même type cellulaire. On peut toutefois distinguer 3 grands groupes: les méthanogènes, les halophiles et les thermoacidophiles.

Les halophiles se développent dans les milieux hyper salins. Leur croissance est souvent associée à celles d'algues halotolérantes comme *Dunaliella*. Les halophiles utilisent plusieurs filières énergétiques (Falb et col., 2008). En condition anaérobie, elles



exploitent l'énergie lumineuse grâce aux bactériorhodopsines et halorhodopsines pour créer un gradient de protons permettant la formation d'ATP via une ATP synthase (Gonzalez *et col.*, 2008).

Les thermoacidophiles auxquels appartient *Pyrococcus horikoshii* se retrouvent dans les environnements à forte activité géothermique comme les sources hydrothermales marines situées sur les dorsales océaniques (fumeurs noirs) ou dans des zones volcaniques (Yellowstone). Certaines espèces comme *Pyrolobus fumarii* se développent jusqu'à 113°C (Stetter, 1999). Les hyperthermophiles sont principalement anaérobies et acidophiles. Ces espèces sont remarquables de par leur grande plasticité métabolique : elles peuvent fonctionner comme hétérotrophes, en utilisant les composés protéiques comme première source de carbone et d'énergie, mais sont aussi chimio autotrophes, en utilisant le soufre élémentaire ( $S^0$ ) ou le  $SO_4^{2-}$  comme accepteurs d'électron pour former du sulfure d'hydrogène (Liu *et col.*, 2012). Au fond des océans, ces organismes sont ainsi à la base d'écosystèmes "chimio synthétiques" qui représentent des oasis de vie dans lesquelles on trouve des eucaryotes (vers, crevettes, crabes et poissons) (Figure 1-21) tous associés via des processus symbiotiques à une biochimie indépendante de la photosynthèse (Chen *et col.*, 2009).

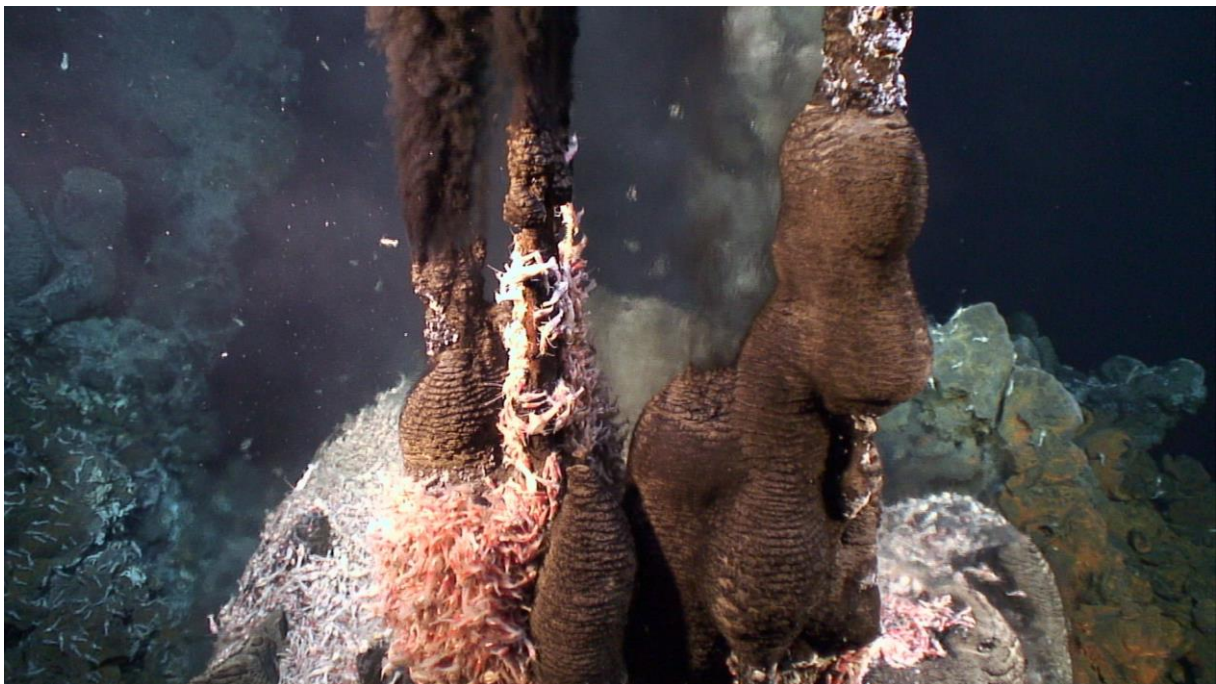


Figure 1-21 : Photo d'un écosystème hydrothermal abyssal (IFREMER/VICTOR 6000)



### 1.6.2 *Pyrococcus horikoshii*, organisme modèle de la thèse

#### 1.6.2.1 Ecologie

*Pyrococcus horikoshii* est une *Euryarchaea* de l'ordre des thermococcales. Ce groupe comprend des organismes qui se développent tous dans des sources hydrothermales marines, en surface ou dans les abysses. Les sources hydrothermales océaniques résultent de remontées d'eau infiltrée dans les fissures du plancher océanique et chauffée du fait d'une activité volcanique. L'activité tectonique est à l'origine de la formation de cheminées appelées "fumeurs noirs" dans lesquels le fluide hydrothermal chargé en minéraux jaillit à une température de 400°C dans une eau à 2°C. La forte pression hydrostatique empêche l'eau de bouillir à ces profondeurs. Les effluents acides sont riches en métaux Fe(II) et Mn(II), et contiennent de fortes concentrations en H<sub>2</sub>S, H<sub>2</sub> et CO<sub>2</sub> (Zierenberg *et col.*, 2000). Ce sont ces composés chimiques réduits qui fournissent les accepteurs et donneurs d'électrons utilisés par les *Thermococcales* pour réaliser une production primaire par chimiosynthèse (Chen *et col.*, 2009). Au contact de l'eau de mer, les minéraux dissous précipitent et forment des édifices creux qui peuvent atteindre plusieurs mètres. Les *Thermococcales* abyssales (*P. horikoshii*, *P. abyssi*, *P. yayanosii*, *T. barophilus*, etc.) proviennent de l'intérieur de ces cheminées. La souche de *Pyrococcus horikoshii* provient de sources hydrothermales situées à 1395m de profondeur au fond de la faille d'Okinawa dans l'Océan pacifique (Gonzalez *et col.*, 1998).

#### 1.6.2.2 Métabolisme

Comme toutes les *Thermococcales*, *Pyrococcus horikoshii* peut fermenter les peptides et les sucres en dioxyde de carbone, hydrogène et acides gras. Si une source de polysulfite est présente les *Thermococcales* dégagent de l'hydrogène sulfuré. Pour le catabolisme du glucose, les *Thermococcales* utilisent un chemin modifié de la voie d'Embden Meyerhoff. Les *Thermococcales* comme les halophiles sont capables d'utiliser le catabolisme des peptides et des acides aminés comme seule source de carbone et d'énergie. Ceci est permis par les amino-transférases qui convertissent les acides aminés en leur 2-oxoacides correspondants par une réaction de transamination avec un autre 2-oxoacide (Kelly *et col.*, 1994; Namboori *et col.*, 2008). Certains acides aminés comme le glutamate, l'aspartate et l'alanine n'ont pas besoin d'être modifiés pour être transaminés et alimenter le métabolisme.

Le matériel cellulaire utilisé pour le travail de thèse a été produit au Laboratoire de Microbiologie des Environnements extrêmes à Brest (IFREMER-Université de Bretagne Occidentale-CNRS). Les cultures ont été réalisées en milieu SME YP à pH 7 (Godfroy *et col.*, 2000), supplémenté en PIPES à 20 mM et en élément soufre à 1 g /L à 90 °C en conditions anaérobies.

### 1.6.2.3 Adaptation thermophile

Les organismes hyperthermophiles sont adaptés aux fortes températures. La croissance optimale de *P. horikoshii* est observée à 98°C, à pression atmosphérique (Gonzalez *et col.*, 1998). Cette propriété passe surtout par une adaptation intrinsèque des composants moléculaires: membranes et protéines essentiellement. Ainsi, les protéines des hyperthermophiles restent stables pendant de longues périodes dans une eau portée à ébullition, conditions dans lesquelles les protéines mésophiles perdent rapidement leur structure repliée. De plus elles montrent un optimum d'activité souvent proche de 100°C alors qu'à des températures en dessous de 50°C, leur activité est fortement ralentie. La structure des protéines est due à des interactions au sein de la chaîne polypeptidique, ou avec d'autres sous unités d'un même complexe et également à des interactions avec le solvant. Les importantes différences de thermostabilité entre enzymes orthologues provenant de mésophiles et d'hyperthermophiles ne sont dues qu'à des substitutions d'acides aminés ou à de petites modifications de la structure. Parmi ces dernières on compte une modification de la taille du cœur hydrophobe de la protéine, du nombre de liaisons hydrogènes ou de paires d'ions, de la taille des boucles et des régions terminales de la protéine ainsi que d'une augmentation des interactions entre sous unités (Jaenicke *et col.*, 1998). C'est la somme de ces "ajustements" qui confère les propriétés thermophiles aux protéines de *Pyrococcus*. Ceci a été montré en comparant les structures cristallographiques d'enzymes orthologues de la famille des malate déshydrogénases (MalDH) (Coquelle *et col.*, 2007). Il faut aussi noter qu'il n'y a pas de règle générale qui s'applique à toutes les protéines car chacune adopte un mode de repliement tridimensionnel qui lui est propre et qui lui confère déjà des propriétés de stabilité. Ainsi une petite protéine globulaire compacte ou une structure quaternaire avec des surfaces de contact étendues seront déjà thermostables et auront moins besoin de modifications pour être adaptées aux conditions extrêmes. Ceci est aussi vrai pour les adaptations halophiles et piezophiles.

Les protéines sont des objets dynamiques qui doivent combiner une rigidité nécessaire au maintien de leur structure 3D avec une flexibilité nécessaire aux mouvements associés à leurs activités biochimiques. Une étude associant la simulation moléculaire à l'analyse de plusieurs structures de MalDH suggère l'existence de "points" chauds dans la structure sur lesquels s'opère des substitutions d'acides aminés destinées à modifier les propriétés dynamiques associées à la fonction (Colletier *et col.*, 2012). Des expériences de diffusion neutronique montrent que, effectivement, l'adaptation thermophile modifie les propriétés dynamiques des protéines. Ceci se traduit par une augmentation de la rigidité (résilience) des protéines hyperthermophiles et le maintien d'une flexibilité aux hautes températures comparable à celle de leurs orthologues mésophile (Tehei *et col.*, 2005). Les mêmes conclusions ont été obtenues en étudiant *in vivo* les propriétés dynamiques du protéome de cellules psychrophiles, mésophiles et thermophiles (Tehei *et col.*, 2004). Ainsi l'accomplissement des fonctions biologiques d'une protéine dans des conditions extrêmes résulte d'un compromis entre stabilité de l'édifice et flexibilité locale nécessaires à la fonction. Ceci explique pourquoi beaucoup d'enzymes thermophiles ne montrent que peu ou pas d'activité à température ambiante alors même que leur structure 3D reste inchangée.

On retiendra que l'adaptation thermophile ne change pas drastiquement la structure des protéines ni leur mode d'action. Ainsi les conclusions d'un travail de génomique structurale et fonctionnelle et de biophysique du mécanisme d'assemblage comme celui présenté dans cette thèse sur les peptidases TET peuvent être transposées aux systèmes mésophiles.

### 1.6.2.4 Intérêt de *Pyrococcus* pour l'étude de la protéolyse et des complexes TET

Le laboratoire ELMA utilise principalement les Archées pour disséquer la fonction de protéolyse. Ce choix est guidé par plusieurs facteurs. Tout d'abord, les Archées présentent un petit génome et ont une position particulière dans l'évolution qui permet plus facilement d'identifier les systèmes et les mécanismes universels qui étaient déjà présents pour faire fonctionner les cellules primitives, voire même LUCA (Gribaldo *et col.*, 2010). De plus, leur proximité évolutive avec les eucaryotes associée au fait que le système PAN-protéasome des Archées est clairement une version moins complexe que le protéasome 26S des eucaryotes devrait permettre de mieux comprendre des aspects

fondamentaux de la biologie de la protéolyse (Maupin-Furlow, 2013). La protéolyse est une fonction qui devient de plus en plus ciblée par l'industrie pharmaceutique. Toutefois, on ne comprend pas bien le rôle de la très grande diversité de peptidases chez les eucaryotes. Leur assignation fonctionnelle est souvent très discutable. Ceci est dû au manque d'informations expérimentales. En effet les peptidases sont des enzymes complexes et il est difficile de "prédire" leurs activités et l'identité de leurs substrats physiologiques sur la simple base de leur séquence, voire même de leur structure 3D. De plus beaucoup d'entre elles sont souvent mal annotées dans les génomes. C'est notamment le cas de la peptidase TET qui est souvent annotée comme une cellulase ou une glucanase. Pour ces raisons il est nécessaire de mener des études intégratives sur les modèles cellulaires "simples" qui combinent des études enzymatiques, des études structurales et des expériences *in vivo*. De ce point de vue les Archées hyperthermophiles offrent de multiples avantages.

Pour ce qui est de la microbiologie, on dispose d'un nombre de génomes séquencés significatif et des outils génétiques ont été récemment développés (Leigh *et col.*, 2011). Reste le problème des cultures anaérobies et le dégagement de H<sub>2</sub>S qui rendent la manipulation des *Thermococcales* délicate. Le choix de ces organismes marins est cependant motivé par le fait que, à travers l'étude de leur fonction protéolytique, on peut aussi aborder des questions sur l'adaptation spécifique de ces organismes aux abysses. En effet, ils sont constamment exposés à des variations importantes de leur environnement (variations de flux hydrothermaux par exemple) qui les exposent à des stress physico-chimiques et nutritionnels fréquents. Ceci les oblige à une grande plasticité métabolique. Dans ces circonstances, la fonction de protéolyse doit être fortement sollicitée soit pour assurer le "contrôle qualité" des protéines via l'élimination ciblée des protéines, soit pour détruire des enzymes devenues inutiles, soit pour fournir de l'énergie à la cellule via la dégradation des protéines (autophagie), soit pour utiliser la matière organique (peptides) de l'environnement comme source d'énergie. A ceci il faut ajouter la régulation, via les peptidases cytosoliques, de fonctions physiologiques ou de processus de défense qui peuvent dépendre de la production de peptides dotés d'activité biologique. Ce dernier aspect de la protéolyse est étudié chez les eucaryotes (transduction de signal, système immunitaire, etc.) mais est peu exploré chez les procaryotes, sauf pour les peptides antimicrobiens. A long terme, l'étude des

aminopeptidases TET au sein du laboratoire ELMA vise aussi à donner des informations dans ce domaine.

L'utilisation des protéines de *Pyrococcus* présente également de grands avantages pour la biochimie structurale des grands assemblages. Tout d'abord leurs propriétés extremophiles permettent de mettre en œuvre des protocoles de purification à température ambiante et, dans le cas d'expression de protéines recombinantes, de réaliser une première étape de clarification des extraits protéiques d'*E. coli* par un choc thermique. Ceci permet de s'affranchir de la tentation d'utiliser des stratégies "faciles" de chromatographie d'affinité utilisant des étiquettes (His-tags etc.). Il est d'ailleurs souvent montré que ces étiquettes perturbent la formation des structures quaternaires. De plus, dans le cas de métallo-enzymes comme TET, cette stratégie est à proscrire. Les propriétés dynamiques des édifices thermophiles les rendent plus aptes à la cristallogénèse. Enfin la thermostabilité et la thermoactivation des complexes hyperthermophiles facilitent aussi la préparation d'échantillons monodispersés et stables dans le temps pour les études structurales en solution comme la diffusion aux petits angles ou l'ultracentrifugation analytique. Il est clair que les études présentées dans cette thèse auraient été beaucoup plus difficiles, voire impossibles à réaliser, avec un système mésophile.

### 1.7 LES PEPTIDASES CYTOSOLIQUES DE *PYROCCOCUS HORIKOSHII*

---

Pour discuter de la place et du rôle des peptidases TET dans la physiologie des cellules, il était important d'avoir une vue d'ensemble de l'équipement protéolytique de *Pyrococcus horikoshii*. Le système de classification MEROPS, déjà présenté dans l'introduction, suggère que la plupart des *Thermococcales* possèdent environ 80 peptidases dans leur protéome. 148 souches d'Archée sont référencées dans la base de données MEROPS, le nombre moyen de peptidases chez les Archées serait de 55, ce nombre varie entre 11 et 148. Il y a peu de différences entre les différentes souches de *Pyrococcus* qui ont été les plus étudiées (*P. furiosus*, *P. abyssi* et *P. horikoshii*).

Ainsi, même dans un organisme procaryote possédant un petit génome on relève une grande diversité de peptidases. Il n'y a que très peu d'analyses qui proposent une vision intégrative de cette diversité. Pour tenter de comprendre le rôle des peptidases ATP-

indépendantes dans la cellule, nous avons tenté de les classer sur la base de leurs fonctions physiologiques présumées et de leurs structures quaternaires. Une identité fonctionnelle est annoncée pour une cinquantaine de peptidases de *Pyrococcus*. Les autres sont présentées comme peptidases putatives ou homologues de peptidases ayant une fonction inconnue. Nous avons complété ces informations avec les données expérimentales (fonctions et structures) de la littérature. Ceci nous a permis de dégager 4 groupes de peptidases qui sont présentés dans la figure 1-22. Pour cette étude, nous nous sommes limités aux peptidases cytosoliques. Ainsi les peptidases associées aux membranes, les peptidases sécrétées et les peptidases autocatalytiques ont été exclues de cette analyse, de même que la vingtaine de peptidases "proposées" mais pour lesquelles il n'y a aucune donnée structurale ou fonctionnelle.

La panoplie des peptidases déployée dans la cellule de *Pyrococcus horikoshii* comprend 35 peptidases réparties dans 20 familles de protéases différentes proposées par le système de classification MEROPS (Rawlings *et col.*, 2012). La majorité (23) sont des membres du clan des métallopeptidases (M) appartenant aux familles: M19, M20, M24, M32, M42, M54, M55 et M82. On trouve ces métallopeptidases dans les 4 grands groupes fonctionnels. Les autres familles sont au nombre de 9, réparties dans 3 clans (A, S et C).

La place des protéases ATP-dépendantes dans les processus de protéolyse est clairement établie. Chez *Pyrococcus* il s'agit du protéasome 20S associé à son complexe activateur AAA ATPase "PAN" (Proteasome Activating Nucleotidase) et de la protéase Lon qui possède un domaine ATPase et est spécialisée dans la dégradation des protéines membranaires. Ces protéases jouent un rôle crucial dans l'élimination des protéines mal repliées ou endommagées et régulent finement beaucoup de processus cellulaires en contrôlant la demi-vie de protéines régulatrices ou d'enzymes clé du métabolisme. En revanche, le rôle des nombreuses peptidases ATP-indépendantes est plus difficile à établir. Ceci concerne en premier chef les peptidases TET qui sont l'objet de la thèse. L'étude de l'identité enzymatique et des informations bibliographiques sur les peptidases de *Pyrococcus* permet de mettre en évidence 4 groupes de fonction.

### 1.7.1.1 Groupe I : Modifications post-traductionnelles

Le groupe I correspond aux fonctions de modification post traductionnelles des protéines. On ne trouve que deux peptidases impliquées dans cette fonction. Il s'agit de

la Methionine aminopeptidase, responsable de l'élimination de la première méthionine des extrémités N-ter. Cette fonction est universelle et contrôle de nombreux aspects de la dynamique intracellulaire des protéines, notamment leur temps de demi-vie (Frottin *et col.*, 2006). La deuxième enzyme est la trilobed peptidase ou PhAAP impliqué dans le déblocage des extrémités N-ter chez les eucaryotes, elle est présentée en section 1.3.5 de ce manuscrit (Tsunasawa, 1998). Chez *Pyrococcus* cette enzyme est auto-compartmentée. La structure a été déterminée en 2014 et indique un mécanisme de sélection de la taille du substrat induit par l'hexamerisation de la protéine (Tsunasawa, 1998). Il s'agit d'une prolyl oligopeptidase (POP) dont la fonction n'est pas clairement établie. Nous pourrions aussi la classer dans le groupe III de notre classification. On peut également placer dans cette catégorie un type d'isopeptidase qui est engagée dans la fonction de modification des protéines par conjugaison. La fonction peptidase est responsable du recyclage d'une étiquette, propre aux Archées, apparentée à l'ubiquitine: les Samp (Humbard *et col.*, 2010). Cette modification post-traductionnelle joue de multiples rôles et il est possible qu'elle soit impliquée dans l'adressage des protéines au système PAN-protéasome (Maupin-Furlow, 2013). Alors que PAN est un homologue du complexe AAA ATPase (sous unités Rpt) qui constitue la base du complexe 19S chez les eucaryotes (Bar-Nun *et col.*, 2012), les protéines JAMM M69 sont possiblement des homologues des protéines Rpn qui font partie intégrante du complexe 19S de régulation des eucaryotes (Cooper *et col.*, 2009).

### 1.7.1.2 Groupe II : Maturations des protéines

Le groupe II correspond à des activités de maturation, spécifique à certaines familles de protéines. FrxA est une endopeptidase de la famille A31 impliquée dans la maturation des précurseurs des hydrogénases à Nickel-Fer NADP-dépendantes (Sun *et col.*, 2010). PrsW est une M82 qui, chez les bactéries, inactive le facteur extra-cytoplasmique anti-sigma RsiW en réalisant les clivages initiaux (Heinrich *et col.*, 2009). Chez les Archées, on ne connaît pas sa fonction. Enfin, *Pyrococcus* possède une pyroglutamyl peptidase de type C15. Chez les Eucaryotes cette enzyme est responsable de la maturation de différentes catégories d'hormones en éliminant le résidu pyroglutamate (pGlu) des extrémités N-terminales (Dando *et col.*, 2003). Encore une fois, on ne connaît pas le rôle physiologique d'une telle enzyme chez les Archées.



### 1.7.1.3 Groupe III : Peptidases résidus ou motifs spécifiques

Le groupe III rassemble des activités peptidases qui ciblent des résidus ou des motifs peptidiques très spécifiques. *Pyrococcus* possède une peptidase appartenant à la vaste famille des papaines/cathepsines. Cette famille (C1) contient principalement des endopeptidases mais elle comprend aussi quelques exopeptidases. Les peptidases C1 contribuent aux activités protéolytiques du système lysosomal et à la présentation des antigènes au système immunitaire des eucaryotes, elle participe également à la dégradation du latex (Vairo Cavalli *et col.*, 2003). Sa présence chez des procaryotes reste mal comprise.

On note aussi la présence de deux di-peptidases de la famille M19, spécialisées dans l'hydrolyse des résidus D-aa. En principe associées aux membranes, ces enzymes stéréospécifiques sont responsables de la résistance à certains antibiotiques comme la vancomycine. L'enzyme archéenne est apparentée à l'enzyme thermophile de *Brevibacillus borstelensis* qui possède une activité d'hydrolyse des acides aminés D et présente également une activité  $\beta$ -lactamase (Baek *et col.*, 2004). Elle est utilisée en biotechnologie pour la production de D-dipeptide ou la synthèse d'agents édulcorants. Deux membres de la famille M24 sont présents et ont été caractérisés chez *Pyrococcus horikoshii*. Ces enzymes sont des prolidases impliquées dans la régulation de la pression osmotique. Elles sont également utilisées pour la détoxification des agents organophosphorés (Theriot *et col.*, 2010a; Theriot *et col.*, 2010b; Zaprasis *et col.*, 2014). Deux peptidases de la famille S9, dotées d'activité spécifique liée aux extrémités N-ter riches en proline (Prolyl aminopeptidases) sont présentes. Elles forment une structure auto compartimentée en trilobe décrite dans le chapitre 1.3.5. Leur rôle physiologique reste obscur (Ishikawa *et col.*, 1998). Le complexe DppA de la famille M55 est un homo-decamer. Chez *Bacillus subtilis* il est impliqué dans l'adaptation aux carences nutritionnelles. Son rôle est de produire des résidus D-Ala pour le métabolisme énergétique (Remaut *et col.*, 2001). Il semble y en avoir deux versions dans *Pyrococcus horikoshii*. Enfin Pab87, autre grand complexe peptidase auto-compartimenté de la famille S12, possède également une activité (carboxypeptidase) spécialisée dans le métabolisme des résidus D-Ala. Il existe également en deux versions chez *Pyrococcus horikoshii*. Etant donné que la paroi des Archées ne possède pas de peptidoglycanes, la

raison de l'existence chez elles de telles machineries spécialisées dans l'hydrolyse des D-amino acides reste donc mystérieuse.

### 1.7.1.4 Groupe IV : Peptidases à large spectre

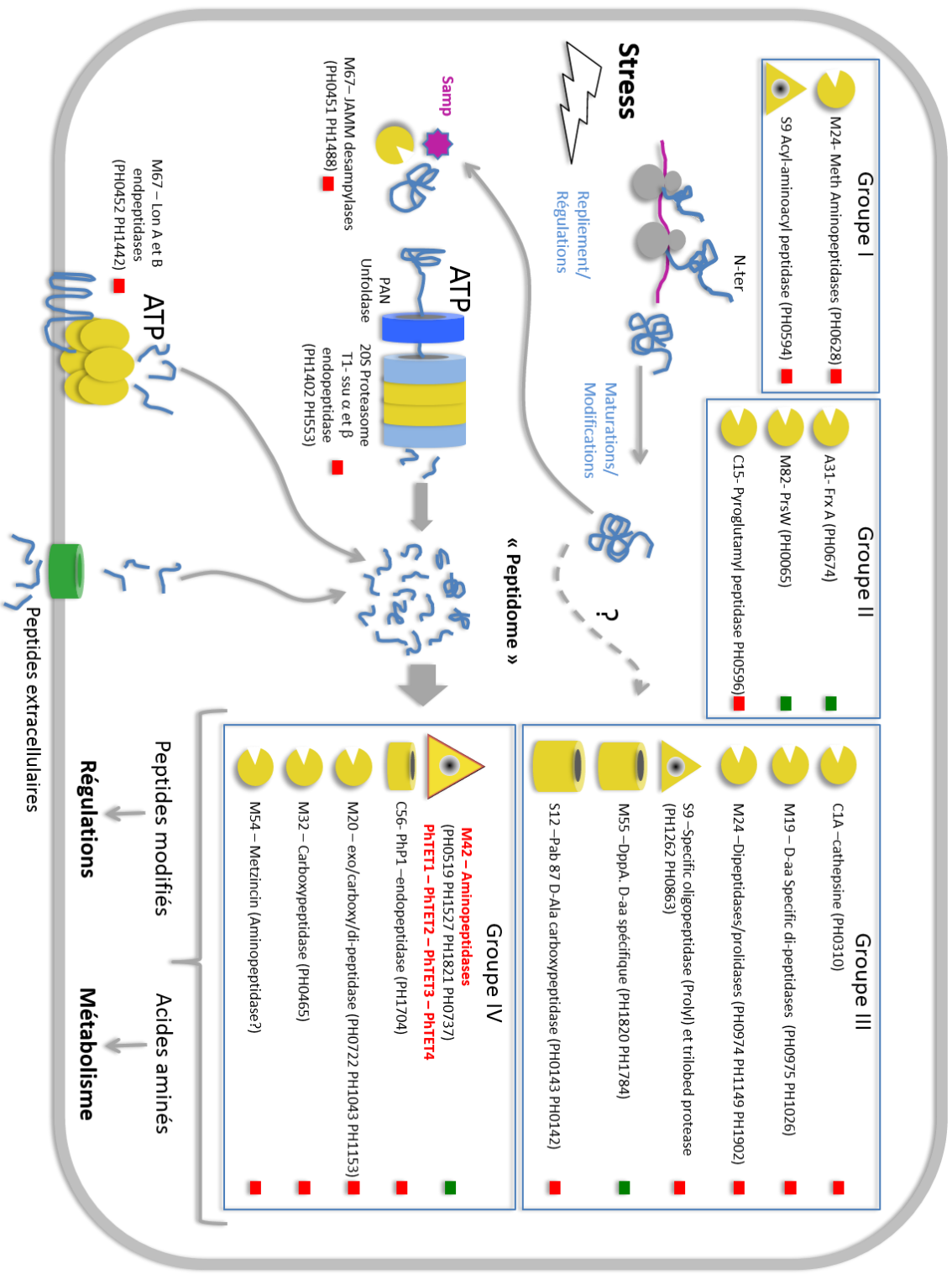
Le groupe IV rassemble des peptidases à large spectre de substrat et donc "a priori" susceptibles de jouer un rôle dans la destruction complète des chaînes polypeptidiques. Il s'agit de trois (jusqu'à 6 si on prend en compte celles qui ne sont pas clairement assignées) metallopeptidases de la famille M20. Les membres de cette famille possèdent des activités très différentes : aminopeptidase, carboxypeptidase ou di-peptidase. Chez *Sulfolobus*, l'une d'elle a été caractérisée comme une carboxypeptidase (Sommaruga *et col.*, 2014). Une caractérisation des différentes peptidases M20 de *Pyrococcus* serait nécessaire. On peut toutefois raisonnablement proposer qu'elles représentent un jeu complet d'activités destinées à la destruction des petits peptides intracellulaires. Le groupe IV comprend également une autre métallopeptidase de la famille M32 qui a, elle, été bien caractérisée chez *P. furiosus*. Elle possède une activité carboxypeptidase à large spectre activée par le cobalt (Cheng *et col.*, 1999). La structure de la protéine de la famille M54 de *Pyrococcus horikoshii*, encore appelée Archaea metzincine, a été déterminée sous une forme inactive (Graef *et col.*, 2012). Une activité aminopeptidase a cependant été proposée pour ses homologues humains (Diaz-Perales *et col.*, 2005). Le groupe IV comporte une endopeptidase hexamérique auto-compartmentée: protéase 1 ou Php1. Cette protéase compte parmi les enzymes les plus thermostables et représente un composé abondant dans le protéome des *Thermococcales*. L'expression de Php1 apparaît être induite en condition de stress. Son activité contre des peptides de moins de 20 résidus a seulement été montrée chez *P. furiosus* (Halio *et col.*, 1997). Des études réalisées dans notre équipe chez *P. horikoshii* n'ont pas permis de montrer une activité. Enfin les 4 protéases TET appartenant à la famille M42, peuvent être rangées dans ce groupe. A l'origine de ce travail, il n'en existait que trois. L'existence d'une quatrième TET a été proposée sur la base des travaux préliminaires présentés dans le chapitre 4 de cette thèse. Les 3 TET de *Pyrococcus* qui ont été caractérisées sont toutes des aminopeptidases strictes. Elles possèdent cependant des préférences marquées pour des acides aminés N-terminaux. Ainsi, PhTET1 hydrolyse préférentiellement des résidus acides et peut se définir comme une Glutamyl-Aspartyl aminopeptidase. TET2 est spécialisée dans les résidus neutres et aliphatiques alors que TET3 hydrolyse

efficacement les résidus basiques et peut être assimilée à une Lysyl aminopeptidase. Les raisons de la complexité structurale et fonctionnelle du système TET de *Pyrococcus* constituent le sujet de fond de cette thèse. Ceci d'autant que collectivement, les autres peptidases du groupe IV sont à même de dégrader des polypeptides complexes.

Dans la figure 1-22 nous avons distingué les peptidases énergie indépendante oligomériques auto-compartmentées de celles qui sont monomériques ou dimériques. Ces peptidases "géantes" ont déjà été présentées dans le chapitre 1.3. On relève que *Pyrococcus horikoshii* en possède 5 dans son cytosol. Il s'agit de DppA, de la D-Ala carboxypeptidase et d'une Prolyl-oligopeptidase spécifique. Ces trois systèmes appartiennent au groupe III spécialisé dans des modifications spécifiques des peptides. Les deux autres, PhP1 et TET, seraient associées au groupe IV et sont "a priori" plutôt impliquées dans des fonctions de destruction des peptides. Tous ces complexes sont impliqués dans la manipulation de peptides de petite taille. Ceci soulève une question qui est au centre de la thèse, à savoir la raison d'une oligomérisation et de la création de structures quaternaires aussi complexes pour des enzymes destinées à manipuler des substrats de petite taille, à la différence des peptidase ATP-dépendantes qui doivent déplier et dégrader des protéines entières.

Sur la figure 1-22 est également indiquée la conservation des différents types d'activité enzymatique vis à vis des trois domaines du vivant. On note que la plupart sont conservées dans les 3 domaines. Ceci ne signifie pas que toutes ces enzymes ont un lien évolutif. Ainsi, l'activité T1 du protéasome est certes conservée entre les bactéries et les eucaryotes, mais les protéines qui portent cette activité : le système HslVU chez les Bactéries et le protéasome chez les Archées et les Eucaryotes n'ont pas de lien évolutif. De même, TET de par son appartenance à la famille M42, fait partie des peptidases qui ne sont en apparence pas présente chez les eucaryotes. Il faut toutefois nuancer ce propos. En effet, comme il sera montré dans le chapitre 4, il existe chez les Eucaryotes un complexe dont la structure est très similaire à TET. Il s'agit aussi d'une aminopeptidase, mais de la famille M18. Les sites actifs des M18 et des M42 sont très proches. Bien que l'analyse phylogénomique permettant d'établir un lien évolutif entre les deux systèmes reste à faire.

De manière générale notre étude montre que la majorité des peptidases intracellulaires des Archées sont destinées à dégrader des peptides de petite taille. Certaines au sein du groupe III correspondent à des spécificités de substrats marquées dirigées contre des peptides cytosoliques. Les rôles physiologiques de ces types d'activités dans une cellule procaryote ne sont pas bien compris et représentent un axe de recherche intéressant. Le groupe IV rassemble des peptidases qui présentent une apparente spécificité de substrat élargie. Collectivement les membres de ce groupe sont donc les plus à même d'assurer la fonction de destruction des polypeptides destinée à produire des acides aminés pour le métabolisme. Les peptidases TET peuvent être rangées dans ce groupe. Au démarrage de mon travail de thèse, les TET apparaissent comme un système de destruction intégré des polypeptides car ils combinaient des spécificités différentes. Cependant mes travaux ont contribué à modifier cette vision. En effet si TET2 et TET3 constituent bien un peptidasome à large spectre d'action, TET1 semble jouer un rôle physiologique différent. Quant à TET4, qui a été découvert dans le cadre de cette thèse, son étude en cours au laboratoire indique qu'il ne fonctionne pas avec les substrats dégradés par les 3 autres TETs et semble donc manipuler des substrats bien spécifiques. La place des TET dans les processus cellulaires reste donc à déterminer.



**Figure 1-22:** Diversité des peptidases cytosoliques chez *Pyrococcus horikoshii*. Les peptidases ont été classées par fonction sur la base de travaux expérimentaux publiés ou en tenant compte de leurs homologues de séquences ou structurales avec des peptidases caractérisées chez d'autres organismes. Les peptidases auto-compartimentées formant des grands assemblages sont représentées par un symbole évocateur de la forme de leur structure quaternaire. Le caractère universel de l'activité peptidase est indiqué par un carré rouge. Celles qui ne sont présentes que chez les procaryotes sont indiquées par un carré vert.

### 1.8 OBJECTIFS ET METHODOLOGIE DE LA THESE

---

Mon sujet de thèse s'inscrit dans la thématique de la protéolyse intracellulaire et de la biophysique des grands assemblages. Il a pour but d'apporter des informations sur le rôle des grandes peptidases TET énergie indépendantes dans le cytosol. Les travaux du laboratoire ELMA avaient permis de décrire la structure de 3 complexes TET présents chez l'Archée *P. horikoshii*. Ils avaient aussi permis de proposer un mode d'action et montré leur identité fonctionnelle en tant qu'aminopeptidases strictes dotées de spécificités différentes. Cependant, le rôle physiologique de ces grands assemblages et leurs liens éventuels avec l'extremophilie restait inconnu. De plus, la publication de travaux de cristallographie et une analyse rapide par alignement de séquence "BLAST" indiquaient la présence, en nombre variable, de peptidases TET dans de nombreux génomes non seulement archéen mais aussi bactériens et eucaryotes. Par ailleurs, des études *in vivo* réalisées au laboratoire indiquaient que, chez l'Archée halophile *H. salinarum*, l'oligomérisation de TET est un processus régulé en fonction des conditions physiologiques. Dans ce contexte, la thèse a pour objectif de répondre aux 4 questions suivantes:

- **Le complexe TET est-il répandu dans les différents types cellulaires?**
- **Pourquoi existe-t-il de multiples versions des peptidases TET dans certains types cellulaire?**
- **Quelle est la signification biologique de l'auto-assemblage de la peptidase TET?**
- **Quels sont les processus d'assemblage de grandes structures quaternaires géométriques?**

Pour répondre à ces questions nous avons utilisé une stratégie intégrative combinant la biologie structurale, la biochimie, la biophysique et la microbiologie.

La comparaison des structures des différentes TET archéennes et bactériennes m'a permis d'identifier des motifs et domaines structuraux conservés entre les différents types de complexes dodécamériques appartenant à la famille enzymatique des peptidases M42. Ce travail est présenté dans le chapitre 4. Il a ainsi été possible de proposer à Simonetta Gribaldo de l'Institut Pasteur un ensemble de critères permettant



d'identifier sans ambiguïté la présence d'édifice TET dans les génomes. Ceci a permis de proposer l'existence d'un quatrième type de peptidase TET chez *P. horikoshii*. Les travaux préliminaires visant à déterminer la structure par cristallographie aux rayons X et l'activité enzymatique de ce nouveau complexe sont présentés dans le chapitre 7.

Le travail de bio-informatique a également permis de proposer une origine de la diversité des TET des *Thermococcales*. Ceci m'a conduit à travailler sur la question des interactions physiques entre les différentes TET de *Pyrococcus*. Nous avons étudié leurs états d'oligomérisation respectifs *in vivo* et leurs propriétés d'hétéro-oligomérisation lors d'expériences de co-expression dans *E. coli* ou de reconstruction de complexe après désassemblage des particules homo-oligomériques. Les propriétés des complexes ainsi formés ont été caractérisées par des mesures enzymatiques et l'analyse de profils de digestion peptidique obtenus par HPLC. Ce travail est présenté dans le chapitre 4.

Alors que le complexe TET est extrêmement stable *in vitro* des expériences *in vivo* montraient l'accumulation d'espèces de bas poids moléculaires de TET dans les extraits cellulaires. Ceci nous a conduits à nous intéresser aux processus d'assemblage des peptidases TET. Ce travail est présenté dans le chapitre 5. Nous avons utilisé une approche de mutagenèse pour ralentir le processus d'assemblage. Il a ainsi été possible d'obtenir par diffraction des rayons X aux petits angles (SAXS) la structure du précurseur dimérique et de caractériser les intermédiaires d'assemblage par ultracentrifugation analytique (AUC) et microscopie électronique. Nous avons également montré la présence du précurseur dimérique dans les cellules de *P. horikoshii*. Ceci, combiné à l'analyse des interfaces entre les sous unités de TET dans les structures cristallographiques, a permis de proposer un modèle pour le processus d'assemblage des peptidases TET, qui apparaît comme un processus très contrôlé. Ce modèle a été validé expérimentalement grâce à une approche innovante présentée dans le chapitre 6. L'étude s'appuie sur le travail de biochimie sur l'hétéro-oligomérisation des TET présenté dans le chapitre 4 pour obtenir des complexes partiellement deutérés. La topologie des sous unités au sein des complexes est révélée par diffraction neutronique aux petits angles (SANS). Ce travail a montré le caractère non-aléatoire du processus d'assemblage du tétraèdre et permet d'expliquer l'efficacité des complexes TET hétéro-oligomériques par rapport aux complexes homo-oligomériques.

## 2. METHODES UTILISEES

Dans ce chapitre, je présenterai les principes des techniques de biochimie puis de biophysique utilisés durant la thèse.

### 2.1 BIOCHIMIE

---

#### 2.1.1 Purification des protéines

##### 2.1.1.1 Expression, lyse et clarification

---

Les protocoles pour la purification des protéines TET utilisées dans cette thèse sont tous basés sur un même modèle utilisé en routine dans le laboratoire pour la purification des grands assemblages natifs provenant d'organismes hyperthermophiles.

La première étape d'expression est réalisée dans l'organisme hôte *Escherichia coli*. Dans le cas de PhTET2, des bactéries compétentes BL21C (DE3)-RIL ont été transformées avec un plasmide pET-41c portant le gène de la protéine PhTET2 ainsi qu'un gène de résistance à la kanamycine selon un protocole classique de transformation. L'expression du gène PhTET2 est régulée par un promoteur lactose sensible à l'IPTG. Les cellules transformées sont sélectionnées sur milieu complet implémenté en kanamycine. Un litre de culture de cellules BL21C (DE3)-RIL transformées, à 37°C dans un milieu LB + chloramphénicol (34 µg/mL) + kanamycine (100 µg/mL) est induit à l'IPTG (0,05 µM) pendant 4h pour atteindre une turbidité à 595 nm de 0,6-0,7. Les bactéries sont ensuite culotées par centrifugation à 4500 rpm (Beckmann J2-21M®, rotor JA-10, 4°C, 15min) et resuspendues dans 50ml de tampon de lyse (50mM TRIS HCl, 150mM NaCl, 0,1% Triton X100, pH = 8,0). Les cellules sont lysées à 4 °C par ultrason (5 fois 30 secondes avec 30 secondes de pause entre chaque impulsion avec une sonde Branson Sonifier 50® à puissance maximale) après ajout d'une solution composée comme suit : lysozyme (0,25 mg/mL) - Dnase I (grade II, Roche 0,05 mg/mL) – Rnase (0,2 mg/mL) – Pefabloc (1 mg/mL) dans 500 µL MgSO<sub>4</sub> 2 M.

L'ensemble des protéines purifiées durant la thèse sont des enzymes provenant de l'Archée hyperthermophile *Pyrococcus horikoshii*. Les protocoles développés dans le laboratoire pour la purification de ces protéines mettent à profit leur grande

thermostabilité à l'étape de clarification. Les lysats cellulaires sont soumis à un choc thermique à 85 °C pendant 15 min afin de provoquer la précipitation de la grande majorité des protéines mésophiles de l'organisme hôte. Après centrifugation du lysat à 17400 g (Beckmann J2-MC®, rotor JA-20, 12000 rpm, 4 °C, 1 h) les protéines surexprimées sont récupérées avec le surnageant.

Par ailleurs, les cultures de cellules de *Pyrococcus horikoshii* utilisées dans les chapitres 4 et 5 ont été cultivées à l'IFREMER de Brest (IFREMER-Université de Bretagne Occidentale-CNRS) en collaboration avec Anne Godfroy. Ces cultures ont été réalisées en milieu SME YP à pH 7 (Theriot *et col.*, 2010b), supplémentées en PIPES à 20 mM et en élément soufre à 1 g/L à 90 °C en conditions anaérobie.

### 2.1.1.2 Chromatographies

---

Lorsque j'ai débuté ma thèse, à la suite de l'étape de clarification, le protocole utilisé impliquait une série de 3 colonnes de chromatographies, une échangeuse d'ions faible (diéthylaminoethyl-cellulose ou DEAE-cellulose), puis une échangeuse d'ions forte (RESOURCE Q) et enfin une colonne de filtration sur gel (SUPEROSE 6).

L'étape de clarification par choc thermique permettant d'éliminer la grande majorité des protéines mésophiles, les purifications réalisées durant la thèse ont été faites en deux étapes de chromatographie seulement, la colonne échangeuse d'ions forte puis la colonne de filtration sur gel.

En plus de son utilisation lors de la purification des protéines recombinantes, la colonne échangeuse d'ions forte RESOURCE Q a été utilisée lors de l'analyse de l'hétéro-oligomère PhTET2:PhTET3 présenté dans les chapitres 4 et 6 de cette thèse. L'éluant utilisé lors des expériences de chromatographie échangeuse d'ion a toujours été le chlorure de sodium NaCl. Les enzymes PhTET2 et PhTET3 ayant des charges électrostatiques de surfaces très différentes, l'aminopeptidase PhTET2 est éluée de la colonne à des conductivités comprises entre 25 et 30 mS/cm, alors que l'aminopeptidase PhTET3 est éluée à des conductivités comprises entre 40 et 45 mS/cm. Le gradient d'élution a donc été adapté tel qu'il est présenté dans le chapitre 4 afin de permettre d'avoir une méthode rapide permettant de cribler les assemblages homo-

oligomériques PhTET2 ou PhTET3 des assemblages hétéro-oligomériques PhTET2:PhTET3.

À la fin de la première étape de chromatographie échangeuse d'ions, le profil d'élution obtenu est complexe et il est difficile de distinguer le ou les pics correspondant à la protéine d'intérêt des autres pics. En effet, l'échantillon déposé contient malgré tout encore de multiples protéines ainsi que des acides nucléiques qui participent au signal d'absorbance observé. Une même protéine étant éluée de manière légèrement différente en fonction de la complexité de l'échantillon de départ, une méthode de criblage permettant d'identifier les fractions contenant la protéine d'intérêt à chaque nouvelle purification est nécessaire. Dans le cas d'aminopeptidases déjà caractérisées, le test enzymatique présenté plus loin dans ce chapitre reste la méthode la plus rapide à mettre en œuvre, elle présente l'avantage de donner un résultat ne présentant quasiment aucune ambiguïté. Dans les cas où le test d'activité n'est pas possible, l'utilisation de l'électrophorèse sur gel de polyacrylamide en conditions dénaturantes (SDS-PAGE) est utilisée afin de différencier les protéines en fonction de leur poids moléculaire. Cependant, comme dit précédemment, les protéines provenant de *P.horikoshii* sont très stables, les dénaturer afin de les analyser sur gel représente donc une difficulté. Pour cela, un protocole de dénaturation spécifique a été développé au laboratoire. Les protéines sont incubées à 90 °C pendant 7 min dans un tampon de dénaturation composé comme suit : 50 mM Tris, 8 M urée, 2 M thiourée, 75 mM DTT, 3 % SDS, 0,05 % bleu de bromophénol à pH = 6.8. Les protéines sont ensuite déposées sur gel et analysées de façon classique.

Lors de l'étude du mutant de PhTET2 présentée dans le chapitre 5 du manuscrit, l'électrophorèse sur gel de polyacrylamide en condition native a été utilisée comme méthode de criblage pour différencier les formes oligomériques de la protéine mutante. Cette méthode permet un criblage rapide tout en limitant les changements de concentration qui auraient influé sur l'état d'oligomérisation. Le protocole utilisé pour ce mutant a été développé pendant la thèse, les conditions de salinité ainsi que le temps de migration ont été optimisés afin d'obtenir la résolution optimale.

Dans le cas de l'analyse des TET de *P. horikoshii in vivo* présentée dans le chapitre 4, le choc thermique ne permet pas l'isolation sélective de protéines. De plus, ces extraits

contiennent très peu de protéines, l'utilisation d'une colonne de type échangeuse d'ion aurait entraîné une trop grande dilution de l'échantillon. Enfin, l'analyse des extraits protéiques S30 n'est pas possible sur colonne de filtration sur gel, l'échantillon contient trop de protéines différentes pour pouvoir les séparer efficacement en fonction de leur taille, sans compter les risques de détérioration de la colonne. Pour réaliser cette étude, c'est la technique de centrifugation à l'équilibre aussi appelée gradient de sucrose qui a été utilisée. Durant l'expérience, un gradient de sucrose est préparé dans un tube adéquat, l'échantillon est déposé délicatement à la surface et l'ensemble est soumis à centrifugation pendant 20 h (voir chapitre 4). Les protéines en solutions sédimentent alors de façon proportionnelle à leur poids moléculaire. Cette technique permet donc de séparer des protéines d'un échantillon donné en condition native et en limitant la dilution de l'échantillon. Après la centrifugation, les gradients sont fractionnés, leur intégrité est contrôlée par la mesure de la conductivité pendant le fractionnement sur un système FPLC. Les différentes fractions sont alors analysées par gel SDS-PAGE ou Western blot.

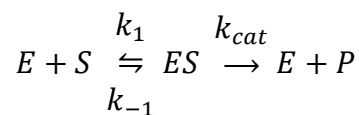
Enfin, dans le chapitre 4, une expérience de chromatographie en phase inverse réalisée sur un système HPLC a été utilisée afin d'analyser l'hydrolyse d'un mélange complexe de peptides par les aminopeptidases TET. Cette technique permet de séparer les molécules en solution en fonction de leur hydrophobicité. Lors de l'expérience réalisée, une colonne  $\mu$ RPC C2/C18 de 0,46×10 cm a été utilisée. Les peptides sont injectés sur la colonne dans un solvant peu hydrophobe (0,065 % acide trifluoroacétique, 2 % acétonitrile), les molécules hydrophobes vont alors être adsorbées sur les billes de silice recouvertes de chaînes alkyles qui composent la résine de la colonne. Le gradient d'élution utilisé est de 0 à 40 % de solvant d'élution (0,05 % acide trifluoroacétique, 80 % acétonitrile). Au fur et à mesure que la concentration en solvant d'élution hydrophobe augmente, les peptides vont être désorbés et donc élués en fonction de leur hydrophobicité.

### 2.1.2 Enzymologie

#### 2.1.2.1 Michaelis-menten

---

Afin de caractériser le fonctionnement d'une enzyme, on réalise généralement des tests enzymatiques durant lesquels la quantité de produit générée par l'enzyme est mesurée en fonction du temps. Le modèle le plus couramment utilisé afin de décrire une réaction enzymatique est le modèle de Michaelis et Menten (Michaelis *et col.*, 2011). Dans ce modèle, une enzyme E fixe un substrat S de manière réversible pour former un complexe intermédiaire ES, une réaction irréversible s'ensuit et mène au relargage du produit P et de l'enzyme libre :



Équation 2-1

La vitesse de la réaction enzymatique  $v$  peut être alors écrite :

$$v = \frac{d[P]}{dt} = k_{cat}[ES]$$

Équation 2-2

L'évolution de la quantité de complexe  $[ES]$  est le résultat de la différence entre le taux de formation du complexe  $[ES]$  et de sa dissociation :

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_{cat}[ES]$$

Équation 2-3

Afin d'intégrer cette équation, il est important de considérer l'hypothèse suivante : excepté durant l'étape transitoire initiale, généralement quelques millisecondes durant lesquelles le substrat et l'enzyme se mélangent,  $[ES]$  est constant tout au long de la cinétique enzymatique jusqu'à ce que le substrat soit presque épuisé, c'est-à-dire que :

$$\frac{d[ES]}{dt} = 0$$

Équation 2-4



Ainsi, selon l'équation 2-3, on a :

$$k_1[E][S] - k_{-1}[ES] - k_{cat}[ES] = 0$$

Équation 2-5

On peut alors exprimer  $[ES]$  en fonction de la quantité d'enzyme totale  $[E_t]$ ,  $[S]$  et les constantes  $k_1$ ,  $k_{-1}$  et  $k_{cat}$  :

$$[ES] = \frac{[E_t][S]}{(k_{-1} + k_{cat})/k_1 + [S]}$$

Équation 2-6

La vitesse initiale de la réaction  $v_0$  peut alors être exprimée en fonction des quantités  $[E_t]$  et  $[S]$  mesurées expérimentalement, c'est l'équation de Michaelis-Menten :

$$v_0 = k_{cat}[ES] = \frac{k_{cat}[E_t][S]}{K_m + [S]}$$

Équation 2-7

Où  $K_m$  est la constante de Michaelis, elle est définie comme suit :

$$K_m = \frac{k_{-1} + k_{cat}}{k_1}$$

Équation 2-8

La vitesse maximale de la réaction  $v_{max}$  est atteinte lorsque le substrat est en excès, dans ce cas  $[E_t] = [ES]$ , donc :

$$v_{max} = k_{cat}[E_t]$$

Équation 2-9

L'équation de Michaelis-Menten s'écrit alors :

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

Équation 2-10

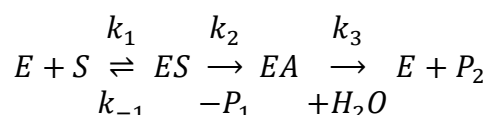
Dans le cas où  $k_{cat}$  est très petit devant  $k_{-1}$ , la constante de Michaelis devient alors :

$$K_m \approx \frac{k_{-1}}{k_1} = K_s$$

Équation 2-11

Où  $K_S$  est la constante de dissociation du complexe  $[ES]$ . Dans ce cas, le  $K_m$  est une mesure de l'affinité de l'enzyme pour son substrat.

Ce modèle peut aussi être appliqué dans le cas de réaction enzymatique impliquant plusieurs étapes distinctes. Dans le cas d'une enzyme michaelienne suivant un processus à deux étapes, c'est le cas de certaines peptidases comme TET, le schéma de réaction est alors le suivant :



Équation 2-12

Dans ce cas, la constante de Michaelis s'écrit :

$$K_m = K_S \frac{k_3}{k_2 + k_3}$$

Équation 2-13

Et  $k_{cat}$  dépend alors des deux constantes  $k_2$  et  $k_3$  :

$$k_{cat} = k_2 k_3 (k_2 + k_3)$$

Équation 2-14

Si  $k_2$  et  $k_3$  ont des valeurs très différentes, alors le  $k_{cat}$  dépendra uniquement de la constante la plus faible.

### 2.1.2.2 Tests d'activité

---

Afin d'analyser les propriétés enzymatiques des aminopeptidases étudiées au laboratoire, des protocoles classiques utilisant des peptides chromogéniques ou fluorogéniques sont souvent utilisés. Ce type de peptide permet d'observer en temps réel l'hydrolyse des substrats en solution puisque la libération du chromophore observée est directement liée à l'activité d'hydrolyse des peptides substrats en solution. Un protocole standard utilisé en routine au laboratoire est présenté dans le chapitre 4.

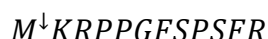
L'analyse de la cinétique de l'hydrolyse du peptide Val-Leu-Lys présentée également dans le chapitre 4 a été réalisée en collaboration avec Jean-Pierre Andrieu à l'IBS, à Grenoble. Le test enzymatique utilisé dans cette expérience est adapté des protocoles standards, l'expérience a été réalisée dans un volume plus grand et une étape de

précipitation des protéines par l'ajout d'acétonitrile glacial, permettant également d'arrêter la réaction, a été ajoutée afin d'analyser uniquement les peptides et acides aminés libres en solution.

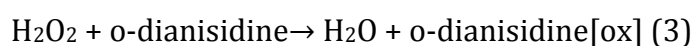
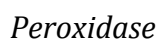
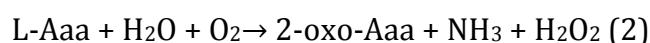
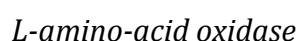
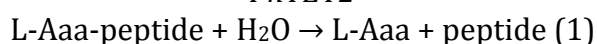
Cependant, certains cas particuliers nécessitent l'élaboration de protocoles originaux pour permettre d'analyser un paramètre en particulier. Lors de l'étude de la relation entre la structure et la fonction de l'aminopeptidase PhTET2, nous voulions tester la différence de prise en charge d'un peptide long entre 2 formes différentes de la peptidase. Dans ce cas précis, seule l'hydrolyse du premier résidu de la chaîne peptidique doit être analysé.

Pour réaliser ce test enzymatique, nous avons adapté un protocole mis au point par Frottin *et col.* (2006) pour l'étude de l'activité de la méthionine aminopeptidase MAP (Frottin *et col.*, 2006). C'est un test indirect qui permet de mesurer la quantité d'acides aminés libres en solution.

La spécificité de substrat de l'aminopeptidase PhTET2 étant forte (voir chapitre 1.5.4), nous avons pu utiliser un peptide substrat dont seul le premier résidu en N-ter peut être hydrolysé :



Ensuite, l'absorption à 440 nm de l'o-dianisidine oxydée est suivie, celle-ci étant proportionnelle à la quantité d'acides aminés libres relargués dans le milieu par l'activité hydrolase de l'aminopeptidase PhTET2 couplée aux activités de la L-amino-acid oxydase et de la peroxydase suivant le schéma suivant :



Connaissant le coefficient d'extinction molaire de l'o-dianisidine ( $10\,580\text{ M}^{-1}\text{ cm}^{-1}$ ), on peut alors déterminer les quantités d'acides aminés libres en solution et donc analyser les cinétiques enzymatiques de PhTET2 sur un peptide long.

Enfin, afin d'analyser l'activité des TET sur un mélange de peptides complexes, l'enzyme a été incubée en présence d'un hydrolysats de caséine. Le mélange de peptides a alors été analysé sur une colonne de chromatographie en phase inverse HPLC comme décrit dans la partie 2.1.1.2. L'analyse des chromatogrammes résultant de cette analyse permet de rendre compte de l'activité générale de l'enzyme sur ce mélange de peptides comme présenté dans le chapitre 4. Dans ce cas, les différents pics observés sur le chromatogramme correspondent aux différents peptides élués pendant l'expérience, la comparaison de l'aire sous ces différents pics observés par rapport à l'expérience contrôle montre ou non l'activité de l'enzyme sur les peptides concernés. Cependant, il est également possible, comme il a été fait avec PhTET1, de purifier les pics de peptides sur lesquels l'enzyme présente une activité et de procéder à leur caractérisation. C'est ainsi que le peptide EGI a pu être défini comme peptide type de l'aminopeptidase PhTET1.

## 2.2 ANALYSE STRUCTURALE

---

### 2.2.1 Dali

---

Le serveur Dali (Holm *et col.*, 2010) est un service en ligne pour la comparaison de structures tridimensionnelles. L'utilisateur fournit un fichier contenant les coordonnées d'une structure d'une protéine afin qu'il soit comparé aux fichiers de structure stockés dans la Protein Data Bank (PDB). Dali utilise un processus de recherche de type « proche en proche ». Autrement dit, lors du processus de recherche, la protéine d'intérêt est comparée aux structures de la PDB, les structures présentant une grande similarité structurale avec la protéine d'intérêt représentent un premier ensemble de structures. Le degré de similarité structurale est calculé et retourné sous la forme d'un facteur nommé Z-score, plus ce facteur est haut, plus la structure est similaire. Ensuite, les fichiers de structures du premier ensemble sont eux-mêmes utilisés en tant que modèle pour chercher à identifier d'autres structures, formant ainsi un deuxième ensemble. Ces structures sont alors comparées à la première protéine d'intérêt fournie par l'utilisateur

pour déterminer le Z-score. Ainsi, les résultats sont retournés sous la forme d'un tableau contenant les structures identifiées en fonction de leur Z-score et donc de leur degré de similarité structurale.

### 2.2.2 Pisa

---

Le serveur PDBe PISA ou 'Protein Interfaces, Surfaces and Assemblies' situé à l'European Bioinformatics Institute ([http://www.ebi.ac.uk/pdbe/prot\\_int/pistart.html](http://www.ebi.ac.uk/pdbe/prot_int/pistart.html)) est un outil en ligne permettant d'explorer les interfaces de macromolécules (Krissinel *et col.*, 2007). L'utilisateur fournit un fichier PDB ou peut en sélectionner un dans la base de données disponible. Cet outil permet alors de déterminer les propriétés structurales et biochimiques des surfaces et des interfaces des macromolécules d'intérêt. Mais il permet également de déterminer toutes les structures quaternaires que pourrait adopter la protéine en fonction des propriétés des résidus à sa surface. Les propriétés structurales et biochimiques, ainsi que des modèles de dissociation peuvent être obtenus pour chacun des complexes hypothétiques proposés. Cet outil permet également de rechercher un type d'interface particulier dans la PDB. En outre, le serveur PISA possède une banque de données d'interfaces pré calculées qui peut être explorée en fonction de nombreux critères tel que le groupe d'espace de la structure, l'aire de la surface d'interaction ou encore en fonction de la présence ou l'absence d'un type d'interaction particulière.

### 2.2.3 Superposition de structures et images

---

L'ensemble des superpositions présentées dans ce manuscrit a été réalisé avec les logiciels COOT (Emsley *et col.*, 2010) et THESEUS (Theobald *et col.*, 2012). COOT (pour Crystallographic Object-Oriented Toolkit) est un logiciel de représentation de structures de macromolécules, il permet l'amélioration et la validation de ces modèles, c'est un logiciel couramment utilisé pour la représentation de protéine en cristallographie aux rayons X. THESEUS est un programme de superposition de structures. Il met à profit une approche par maximum de vraisemblance. Le résultat obtenu avec THESEUS prend mieux en compte la structure dans son ensemble que ne peut le faire COOT, à travers les

superpositions par moindre carré ou en considérant la structure secondaire (Secondary Structure Matching).

L'ensemble des images présentées dans ce manuscrit a été généré à l'aide du logiciel PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.). C'est un logiciel puissant et polyvalent pour la représentation tridimensionnelle de macromolécules.

### 2.3 BIOPHYSIQUE

---

#### 2.3.1 Diffusion aux petits angles

La technique de diffusion de rayons X (SAXS) ou de neutrons (SANS) aux petits angles permet d'obtenir des informations structurales à basse résolution sur des molécules en solution (Jaenicke *et col.*, 1998; Svergun *et col.*, 2013). Ces deux techniques reposent sur le fait que, tout comme la lumière visible, les rayons X et les neutrons peuvent être diffusés lorsqu'ils interagissent avec la matière. Les rayons X vont être sensibles à la densité électronique d'une particule, alors que les neutrons, eux, vont interagir avec les noyaux atomiques. Les distances sondées par la diffusion aux petits angles sont plus grandes que les distances interatomiques, ce sont des distances de l'ordre du nanomètre. À titre de comparaison, la diffraction de rayons X permet de sonder des distances de l'ordre de l'Angstrom, et dans le cas de la diffusion de la lumière, les distances sondées sont de l'ordre du micromètre.

Lors d'une expérience de diffusion aux petits angles (SAS), une solution tampon contenant les macromolécules est placée dans un faisceau de rayons X ou de neutrons et la diffusion aux alentours du faisceau incident qui en résulte est alors enregistrée par un détecteur. Chaque particule de la solution va diffuser et c'est l'ensemble de ces diffusions qui est observée par un détecteur placé en aval de l'échantillon. L'image obtenue a pour centre le faisceau incident et, parce que les molécules en solutions ont des orientations aléatoires, leurs intensités diffusées forment un cliché avec une symétrie concentrique au détecteur par rapport au faisceau incident. Les différents cercles concentriques par rapport au centre de l'image représentent donc les intensités de diffusion de l'échantillon en fonction de l'angle de diffusion. En intégrant les intensités de ces différents cercles concentriques, on obtient la courbe de diffusion unidimensionnelle de



l'échantillon. Sur cette courbe, aux plus petits angles, ce sont essentiellement les structures grossières de la protéine qui dominent le signal, c'est-à-dire sa forme et son extension globales. Les variations observées à des angles plus grands correspondent à des propriétés structurales de courtes distances au sein de la protéine.

Pour qu'une expérience de diffusion aux petits angles fonctionne bien, il est nécessaire que les molécules en solution soient assez diluées pour se soustraire des interactions inter-particulaires, on parle alors d'une solution « idéale ». Si l'échantillon en solution est monodisperse, c'est-à-dire qu'il n'est composé que d'un seul type de particule, des informations structurales sur la particule en solution peuvent être obtenues. Ces informations incluent la masse moléculaire de la particule  $M$ , son rayon de giration  $R_G$  ainsi qu'une reconstruction de l'enveloppe de la particule à une résolution de l'ordre du nanomètre. Dans le cas d'un échantillon polydisperse dans une solution idéale, des valeurs moyennes de  $R_G$  et  $M$  sont obtenues.

### 2.3.1.1 Théorie

---

Dans le cas de la diffusion aux petits angles, on suppose que le faisceau incident de neutrons ou de rayons X n'échange pas d'énergie avec l'échantillon, on parle alors de diffusion élastique. Avec cette hypothèse, les faisceaux incidents et diffusés ont la même longueur d'onde  $\lambda$ . Si  $\vec{k}$  et  $\vec{k}_0$  sont respectivement les vecteurs d'onde des faisceaux diffusés et incidents, et que le vecteur de diffusion  $\vec{q} = \vec{k} - \vec{k}_0$ , alors :

$$q = |\vec{q}| = \frac{4\pi}{\lambda} \sin(\theta)$$

Équation 2-15

Où  $2\theta$  est l'angle de diffusion mesuré. Durant l'expérience, la longueur d'onde  $\lambda$  du faisceau incident est fixe. La valeur de  $q$  ne dépend donc que de l'angle  $\theta$ , les variations d'angle de diffusion sont représentées par  $q$  (en  $\text{\AA}^{-1}$ ) sur les figures présentées dans la suite.

L'intensité diffusée d'une particule dans une solution idéale se calcule à l'aide de la transformée de Fourier de la densité de longueur de diffusion de ses électrons (SAXS) ou noyaux (SANS), en tenant compte des phases relatives du rayonnement diffusé au sein de la particule :

$$i_n(q) = \left| \int \Delta\rho e^{iqr} dV \right|^2$$

Équation 2-16

La densité de longueurs de diffusion  $\Delta\rho$  se calcule avec les longueurs de diffusion (de neutrons ou rayons X)  $b_i$  et le volume  $V_i$  (à la position  $r$ ) que chaque atome occupe au sein d'une particule. Comme la particule est immergée dans une solution de densité de longueurs de diffusion  $\rho_s$  il faut en retirer celle-ci :

$$\Delta\rho = \frac{b_i}{V_i} - \rho_s$$

Équation 2-17

Dans une solution idéale les particules sont orientées de manière isotrope et l'intensité de diffusion a une symétrie sphérique autour du rayon incident. L'intensité de diffusion de l'échantillon en fonction de l'angle est notée  $I(q)$ , elle représente la somme des intensités diffusées par chaque molécule de la solution, moyennée sur toutes les orientations possibles des molécules (notée  $\langle \dots \rangle$  dans l'équation 2-18) :

$$I(q) = \sum_n N_n \langle i_n(q) \rangle$$

Équation 2-18

Avec  $N_n$  le nombre de particules de type  $n$  et  $i_n(q)$  l'intensité de diffusion par particule de type  $n$ .

### 2.3.1.2 Interprétation des données

---

#### 2.3.1.2.1 POIDS MOLECULAIRE M

L'intensité diffusée à l'angle zéro (faisceau direct)  $I(0)$  est proportionnelle à la masse moléculaire  $M$  et à la concentration  $c$  des particules en solution, à condition que les particules aient la même densité de longueurs de diffusion (i.e. la composition chimique moyenne des particules est identique) (Jacrot *et col.*, 1981) (voir équation 2-19).

$$I(0) \propto \sum_n c_n M_n$$

Équation 2-19

L'intensité à  $q = 0$  ne peut être mesurée directement, cette zone étant masquée par le dispositif d'arrêt de faisceau,  $I(0)$  doit par conséquent être extrapolée depuis les données expérimentales (voir section suivante). Si le nombre de différentes particules est faible, et que leur masse moléculaire  $M_n$  ou leur concentration  $c_n$  respectives sont connus, les paramètres inconnus  $M_n$  ou  $c_n$  peuvent être obtenus à partir du  $I(0)$  en utilisant l'équation 2-19.

### 2.3.1.2.2 RAYON DE GIRATION

L'approximation de Guinier (Guinier, 1939) permet d'obtenir une approximation des intensités  $I(q)$  aux très petits angles, autrement dit pour  $qR_G \leq 1$  on a :

$$I(q) = I(0) \exp\left(-\frac{q^2 R_G^2}{3}\right)$$

Équation 2-20

Où  $R_G$  est appelé le rayon de giration. Si l'on trace  $\ln(I)$  en fonction de  $q^2$  la pente à de très petits angles correspondrait alors à :  $-R_G^2/3$  et on peut ainsi déduire le rayon de giration. En extrapolant la droite à l'ordonnée, on peut déduire  $I(0)$ .

Le rayon de giration  $R_G$  décrit la distribution de la densité de diffusion (par rapport au tampon) autour du centre de densité de diffusion de la particule. Les objets sphériques auront des plus petits rayons de giration pour une particule d'une masse donnée.

Si la solution est monodisperse, le rayon de giration obtenu avec l'équation 2-20 peut être associé à la forme et à la taille de la particule en solution. Dans le cas de système polydisperse, la valeur de  $R_G$  est influencée par l'ensemble des particules en solution et sa valeur sera alors une moyenne.

Une autre méthode pour déduire le rayon de giration consiste à utiliser la fonction de distribution des paires  $P(r)$ . Cette fonction correspond en quelque sorte à un histogramme des distances interatomiques. C'est la transformée de Fourier de la courbe des intensités de diffusion  $I(q)$  d'un échantillon. La fonction  $P(r)$  permet de déduire des valeurs de  $I(0)$  et de  $R_G$  en prenant en compte l'ensemble des valeurs collectées et non

pas seulement les valeurs à faible  $q$ . Elle permet également d'estimer la valeur de  $D_{max}$ , la distance maximale entre deux points dans une même particule. Enfin, la forme de la courbe de  $P(r)$  est caractéristique de la forme globale de la particule, cette fonction permet ainsi un contrôle rapide de qualité.

### 2.3.1.2.3 COURBE DE DIFFUSION THEORIQUE

Etant donné que l'intensité de diffusion observée pendant une expérience de SAS est la moyenne sur toutes les orientations des transformées de Fourier des intensités de longueur de diffusion  $\rho_n$ , ou des densités électroniques  $\rho_e$  des particules en solutions, il est possible de calculer les courbes de diffusion théorique si  $\rho_n$  ou  $\rho_e$  sont connus. Les programmes CRY SOL (Svergun *et col.*, 1995) (pour les rayons X) et CRYSON (Svergun *et col.*, 1998) (pour les neutrons) font partie du paquet de programme ATSAS (<http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html>). Ils permettent de calculer les courbes de diffusion théorique de protéines à partir des fichiers PDB correspondants. Les courbes ainsi obtenues sont d'une grande aide pour l'interprétation des profils de diffusion expérimentaux, ou encore pour vérifier si la structure du fichier PDB correspond à la structure de la protéine en solution.

### 2.3.1.3 Variation de contraste en SANS

---

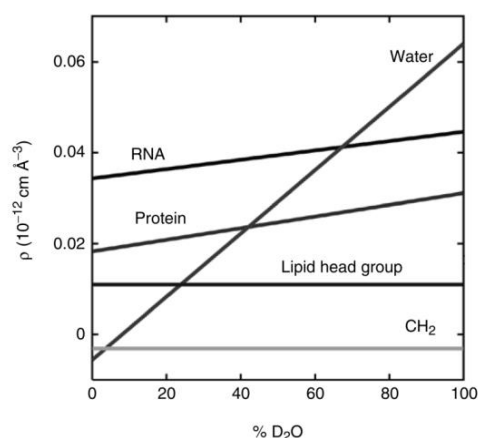
L'avantage de la diffusion de neutrons est la possibilité de distinguer la diffusion des différents constituants d'un système complexe, à condition que les constituants aient des densités de diffusion des neutrons différentes. Dans le cas de complexes multi-protéiques, la deutération sélective d'une protéine du complexe permet d'altérer de façon sélective la densité de diffusion des neutrons de ce partenaire. En conséquence, en variant le niveau de deutération du solvant (par mélange H<sub>2</sub>O-D<sub>2</sub>O), on peut faire varier le contraste de diffusion des neutrons de chaque partenaire. Dans le cas où la densité de diffusion du solvant devient égale à celle du partenaire marqué, celui-ci devient « invisible » en moyenne, il est confondu avec le solvant et les intensités de diffusion alors mesurées ne seront dues qu'au partenaire dont la densité de diffusion diffère de celle du solvant.

Il existe une grande différence entre la longueur de diffusion du deutérium ( $b_D = 0,667 \times 10^{-12}$  cm) et celle de l'hydrogène ( $b_H = -0,374 \times 10^{-12}$  cm), la dernière étant même

négative. C'est cette particularité de l'hydrogène, constituant majeur des protéines, qui fait toute la puissance de cette technique pour les études structurales de molécules biologiques. Cette caractéristique est utilisée en SANS lors de la préparation des mélanges de solvants isotopiques comme le mélange H<sub>2</sub>O-D<sub>2</sub>O. L'eau légère (H<sub>2</sub>O) a une densité de longueur de diffusion de  $-0,562 \times 10^{-14} \text{ cm } \text{\AA}^{-3}$ , alors que celle de l'eau lourde (D<sub>2</sub>O) est de  $6,404 \times 10^{-14} \text{ cm } \text{\AA}^{-3}$ . Donc, la densité de longueur de diffusion du mélange est une fonction linéaire du pourcentage de D<sub>2</sub>O et atteint zéro pour 8 % v/v de D<sub>2</sub>O. A cause des échanges de proton possibles, la densité de longueur de diffusion d'une protéine varie lorsqu'elle est en présence du solvant contenant du D<sub>2</sub>O et est de plus fonction du ratio H<sub>2</sub>O-D<sub>2</sub>O.

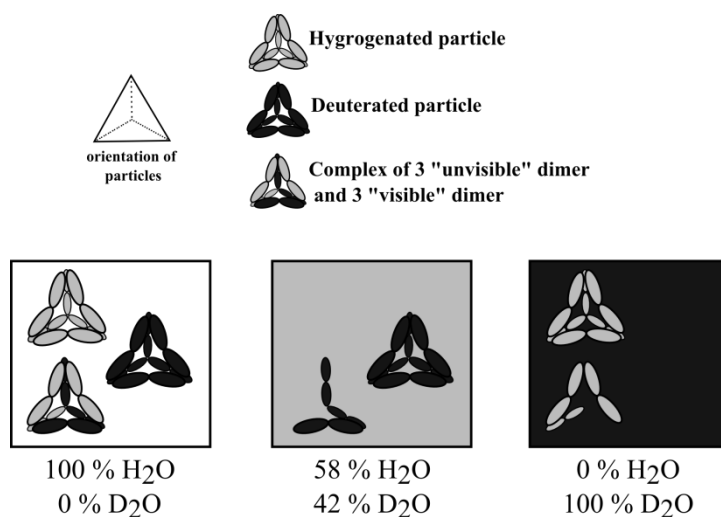
Il existe également une variation de la longueur de diffusion entre les différents atomes constituant les macromolécules biologiques. De ce fait, les grandes classes de molécules biologiques ont des densités de longueur de diffusions différentes. Les acides nucléiques, ADN et ARN, ont une longueur de diffusion plus grande que celle des protéines, qui ont elles-mêmes une longueur de diffusion plus grande que celle des lipides (Harroun *et col.*, 2006). Ceci est représenté graphiquement sur la figure 2-1, on peut y voir les densités de longueur de diffusion de l'ARN, des protéines et des lipides, représentés en fonctions de la concentration en D<sub>2</sub>O. Les points où la droite de l'eau et celles des différentes molécules se croisent est appelé point de correspondance, ils correspondent aux pourcentages de D<sub>2</sub>O nécessaires pour atteindre un contraste nul pour ces molécules.

On remarque qu'à 0% v/v D<sub>2</sub>O, la densité de longueurs de diffusion des protéines est supérieure à celle de l'eau ; à environ 42 % v/v D<sub>2</sub>O, le point de correspondance est atteint et le rapport s'inverse.



**Figure 2-1 : Densité de longueur de diffusion des principaux types de macromolécules biologiques en fonction de la concentration de D<sub>2</sub>O du solvant. A 100 % v/v D<sub>2</sub>O, il est supposé que tous les hydrogènes d'une molécule sont échangés. Les valeurs représentées sont des valeurs moyennes (tirée de Dutoit *et col.* (2012)).**

Dans le cas d'un complexe composé de plusieurs protéines, on peut modifier sélectivement la longueur de diffusion d'un des partenaires par le procédé de deutération. En faisant ensuite varier la densité de longueurs de diffusion du solvant on mesure sélectivement la diffusion d'un partenaire spécifique. À 0% v/v D<sub>2</sub>O, le partenaire hydrogéné, comme le partenaire deutéré, a une densité de longueur de diffusion différente de celle du solvant. À environ 42% v/v D<sub>2</sub>O, la densité de longueur de diffusion du partenaire hydrogéné est identique à celle de l'eau, seul le partenaire deutéré contribue au signal. À 100% v/v D<sub>2</sub>O, c'est la situation inverse, la particule deutérée (à un certain degré) est confondue avec le solvant, seule la particule hydrogénée contribue au signal. Ces phénomènes sont illustrés sur la figure 2-2.



**Figure 2-2 : Illustration schématique de la variation de contraste appliquée à un complexe protéique dont les monomères ont la même conformation. À 0 % v/v D<sub>2</sub>O les deux partenaires sont visibles, à 42 % v/v D<sub>2</sub>O seul le partenaire deutéré est visible et enfin à 100 % v/v D<sub>2</sub>O seul le partenaire hydrogéné.**

### 2.3.2 Ultracentrifugation analytique

Lors d'une expérience d'ultracentrifugation analytique (AUC), c'est la sédimentation des particules en solution qui est analysée. Cette technique permet de caractériser les interactions des macromolécules en solution en observant leur répartition en fonction du temps lorsqu'elles sont soumises à un champ de force centrifuge très important (jusqu'à 300 000 g). On peut ainsi quantifier l'homogénéité des préparations, la forme globale, la masse ainsi que la stœchiométrie des macromolécules. C'est un outil puissant permettant de lever l'ambiguïté sur la nature des protéines en solution.



### 2.3.2.1 Dispositif expérimental

---

Les solutions de protéines sont placées dans des cellules de centrifugation, dites à double canaux, qui sont constituées de deux chambres. Dans une des chambres est injecté l'échantillon à analyser, autrement dit la protéine dans son tampon, et dans la seconde le tampon seul pour référence. On peut suivre la sédimentation des protéines la mesure de l'absorbance à 280nm durant la centrifugation. On peut aussi la suivre en mesurant la différence d'index de réfraction entre les deux compartiments, qui est directement liée à la différence de concentration en protéine et peut être observé par des changements dans le profil d'interférence, et ce, à des temps régulier. Ainsi on détermine la distribution radiale des macromolécules dans la cellule en fonction du temps et donc la vitesse de sédimentation des molécules en solution.

### 2.3.2.2 La sédimentation

---

La sédimentation d'une particule soumise à une accélération centrifuge  $\omega^2 r$  et qui sedimente à la vitesse radiale  $u$  peut être décrite par son coefficient de sédimentation  $s$ , qui peut être calculé par l'équation de Svedberg (Schachman, 1957) :

$$s = \frac{u}{\omega^2 r} = \frac{M(1 - \bar{v}\rho)}{N_A f}$$

Équation 2-21

Où  $M$  est la masse molaire de la particule qui sédimente,  $\bar{v}$  est son volume partiel spécifique,  $\rho$  la densité du solvant,  $f$  le coefficient de friction et  $N_A$  le nombre d'Avogadro. Donc, La vitesse de sédimentation d'une protéine est caractérisée par son coefficient de sédimentation  $s$ , il nous informe sur la taille de la protéine. En effet, à partir du coefficient de sédimentation, et en faisant l'hypothèse que toutes les protéines ont la même densité, on peut en déduire leur poids moléculaire. Le coefficient de friction d'une sphère parfaite  $f_0$  peut-être dérivée de la loi de Stoke :

$$f_0 = 6\pi\eta R_0$$

Équation 2-22

Dans cette équation,  $\eta$  représente la viscosité de la solution et  $R_0$  représente le rayon de la sphère. Le coefficient de friction  $f$  des particules ayant une autre forme est relié au rayon hydrodynamique  $R_H$  :

$$f = 6\pi\eta R_H$$

Équation 2-23

Le coefficient de friction  $f/f_0$  d'une particule traduit la déviation de sa forme par rapport à une sphère. En insérant l'expression de  $f_0$  dans l'équation de Svedberg (Équation 2-21), le coefficient de sédimentation maximal  $s_{max}$  pour une particule de masse moléculaire donnée peut être calculé. En comparant la valeur  $s$  expérimentale à la valeur  $s_{max}$ , le coefficient de friction peut être déterminé et des informations sur la forme globale de la molécule peuvent alors être déduites.

### 2.3.2.3 Traitement des données

---

L'équation de Lamm (Lamm, 1929) décrit la distribution des concentrations en fonction du coefficient de sédimentation des protéines, de l'accélération centrifuge et du coefficient de diffusion  $D$ , ce dernier étant dépendant des conditions expérimentales et du coefficient de friction des protéines. Durant ma thèse, j'ai analysé les résultats avec le logiciel SEDFIT (Schuck, 2000). Ce logiciel permet d'ajuster, à partir de l'équation de Lamm (Lamm, 1929), la meilleure combinaison possible entre coefficient de sédimentation  $s$  et celui de diffusion  $D$ , via le rapport de coefficient de friction  $f/f_0$ , à partir des données expérimentales. On obtient ainsi la distribution des coefficients de sédimentation  $c(s)$  dans l'échantillon analysé.

Une fois les coefficients de sédimentation expérimentaux déterminés, ils peuvent être comparés aux valeurs théoriques calculées à partir de l'équation de Svedberg (Équation 2-21). Une des applications possibles est la détermination de l'état oligomérique d'une protéine puisque des oligomères différents auront des masses molaires différentes.

Afin de comparer les coefficients de sédimentation obtenus lors d'expériences dans des solvants différents, un coefficient de sédimentation normalisé  $s_{20,w}$  est calculé. Ce coefficient de sédimentation corrigé reflète la façon dont sédimente la particule dans l'eau à 20 °C. La relation entre le coefficient de sédimentation expérimental  $s$  et  $s_{20,w}$  est décrite dans l'équation 2-24.

$$s_{20,w} = s \frac{(1 - \rho_{20,w} \bar{v})}{(1 - \rho \bar{v})} \frac{\eta}{\eta_{20,w}}$$

Équation 2-24

Dans cette équation,  $\rho$  et  $\eta$  sont respectivement la densité et la viscosité du solvant utilisé durant l'expérience,  $\rho_{20,w}$  et  $\eta_{20,w}$  sont ceux de l'eau à 20 °C.

### 3.2.1 Cristallisation des protéines

Durant la thèse, j'ai pu m'initier à la cristallographie. Des essais de cristallisation ont été réalisés avec les formes naturelles et mutantes de la protéine PhTET2 ainsi qu'avec PhTET4 (chapitre 7).

Un cristal peut être décrit en définissant une maille, contenant une ou plusieurs copies de la protéine, placée aux nœuds d'un réseau tridimensionnel. Les protéines sont donc organisées et rangées de façon très définie dans le cristal. La qualité d'un cristal dépend de la qualité de l'arrangement des molécules dans ce même cristal.

Afin de déterminer les conditions de cristallisation d'une protéine, l'utilisation d'une méthode de criblage automatisé permet de pouvoir tester un très grand nombre de conditions en peu de temps. À l'IBS, nous avons accès à la plateforme de criblage automatisé HTXlab de l'EMBL à Grenoble, disposant d'un robot de cristallisation nanogoutte. À partir de 80  $\mu$ l de solution de protéine, il est possible de tester 576 conditions de cristallisation différentes. Plusieurs types de cribles standards sont possibles. Un système automatisé de visualisation permet de comparer rapidement les différentes conditions et d'identifier parmi celles-ci les conditions favorables à la cristallisation de la protéine d'intérêt. Ces conditions seront ensuite explorées manuellement au laboratoire.

Pour obtenir un cristal, la méthode la plus commune est la diffusion de vapeur. Pour cette méthode, la solution de protéine à haute concentration est mélangée avec une solution contenant des précipitants, des sels, le tampon ainsi que d'autres molécules, qui est appelée solution précipitante. Le rapport entre la solution de protéine et la solution précipitante est en générale de 1 :1 en première approche, celui-ci peut changer en fonctions des conditions d'optimisation. La goutte obtenue est placée dans un récipient clos en présence d'un excès de solution mère, sans être en contact direct. Par effet de diffusion de vapeur, la composition de la solution dans laquelle sont les protéines est modifiée graduellement et va permettre de réduire la solubilité des molécules en solution. De petits ensembles de protéines vont se former aléatoirement, ils sont

maintenus par des interactions faibles entre les molécules de protéines. Au départ, le rapport de la surface au volume est grand et seul un faible nombre d'interactions stabilisantes sont formées. Au-delà d'une taille critique, le nombre d'interactions devient suffisant pour que l'ajout de molécule soit énergiquement favorable. Une fois cette taille critique dépassée, l'agrégat devient un noyau de cristallisation qui va continuer de croître. La taille critique que doit dépasser un agrégat pour devenir un noyau de cristallisation, et la vitesse de cristallisation, dépendent de la force des interactions entre les molécules et du niveau de sursaturation de la solution. Si la vitesse de cristallisation est trop grande, la protéine va tendre à s'agréger. En modifiant la composition de la solution de cristallisation, la solubilité de la protéine et les interactions entre les différentes molécules peuvent être influencées.



### 3. IDENTIFICATION DE MOTIFS STRUCTURAUX CONSERVES POUR L'ASSIGNATION DES PEPTIDASES TET DANS LES GENOMES ARCHEE

Comme j'ai pu le souligner dans l'introduction, je me suis rendu compte au cours de la thèse que l'assignation des peptidases est souvent erronée, en particulier dans les cas des peptidases M42 qui forment des grands assemblages en forme de tétraèdre (TET). L'aminopeptidase PhTET1 par exemple était assignée en tant qu'endoglucanase avant que sa caractérisation ne soit réalisée (voir 1.5). Cet exemple reflète la méconnaissance des caractéristiques propres aux complexes TET. Les travaux présentés dans la suite de ce chapitre ont donc pour objectif de définir des marqueurs structuraux fiables afin d'identifier sans ambiguïté les aminopeptidases TET dans les génomes.

#### 3.1 INTRODUCTION

---

Après la découverte des complexes peptidases TET chez l'Archée halophile *Haloarcula marismortui*, plusieurs complexes TET ont été caractérisés chez *Pyrococcus horikoshii* (voir 1.5). Ces différents complexes enzymatiques présentent une grande homologie structurale, cependant, le pourcentage d'identité de séquence entre ces TET n'est pas très élevé. Autrement dit, certaines structures secondaires sont disposées de façon très similaire dans les différentes structures de TET malgré le fait que leurs séquences diffèrent. En effet, PhTET1 montre environ 37 % d'identité (55 % de similarité) pour PhTET2 et PhTET3, et ces deux dernières ont 48 % d'identité (70 % de similarité) entre elles.

Récemment, plusieurs structures cristallographiques d'aminopeptidases présentant une grande similarité structurale avec les TET caractérisées chez *P. horikoshii* ont été résolues chez des bactéries et également chez les eucaryotes (Figure 3-1) (Kim *et col.*, 2010; Marguet *et col.*, 2013). Ces peptidases ne présentent que très peu d'identité de séquences entre elles ou avec les PhTET, ce qui rend très difficile leur identification.

Ces données suggèrent que les machines tétraédriques TET seraient présentes dans beaucoup de types cellulaires et dans les 3 domaines du vivant mais que leur identification sur la simple base de recherche d'homologie de séquence est difficile. Pourtant, il est important de pouvoir réaliser des études génomiques et phylogénomiques afin de comprendre le ou les rôles physiologiques des peptidases TET et les raisons de l'existence de plusieurs versions du complexe TET dans une même cellule.

L'existence de ces données structurales et l'ambiguïté quant à l'appartenance de ces complexes au groupe des TET posent plusieurs questions :

- 1) Comment identifier une aminopeptidase TET sur la base de données structurales ?
- 2) Peut-on extraire des éléments structuraux suffisamment conservés entre les peptidases TET et s'en servir pour identifier sans erreur les peptidases TET présentes dans les différents génomes?

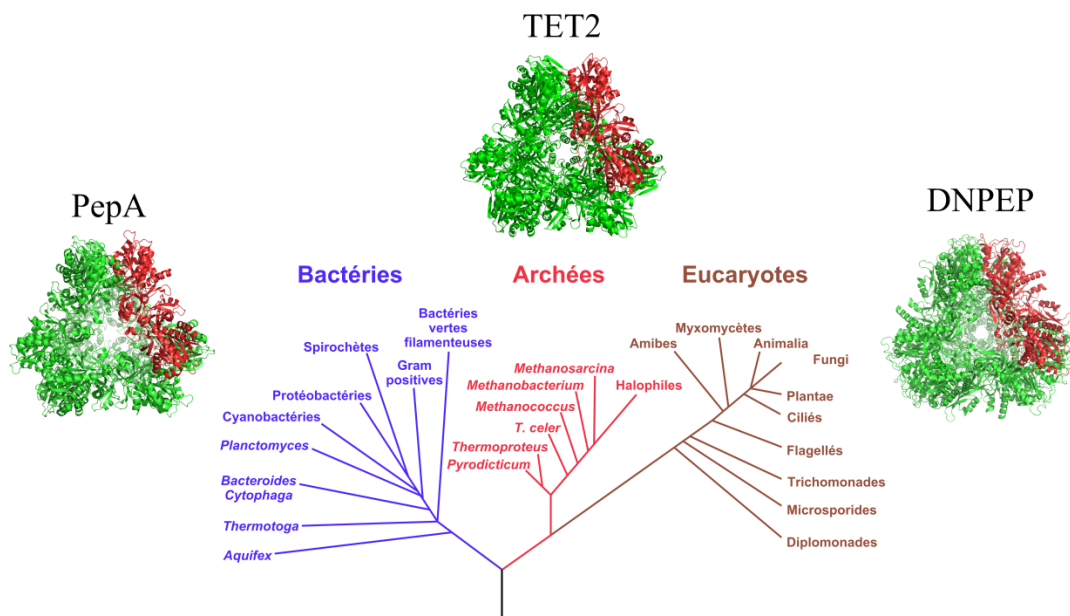


Figure 3-1 : Des TET existeraient dans les trois domaines du vivant. Arbre phylogénétique du vivant (d'après Woese et al. 1990) et les structures cristallographiques de peptidases identifiées comme TET potentielles dans chacun des domaines du vivant : les aminopeptidases bactérienne PepA (Code PDB : 3KL9), archéenne TET2 (Code PDB : 1Y0R) et eucaryotes DNPEP (Code PDB : 4DY0).

Afin de répondre à ces questions, la première étape a donc été d'identifier des marqueurs structuraux pour définir ce qu'est un assemblage de type TET.



L'assemblage d'une particule telle que TET dépend de la présence de domaines structuraux, de la nature des interfaces permettant la formation du complexe quaternaire et de motifs permettant la fonction peptidase. Dans ce chapitre, nous allons examiner ces différents aspects.

### 3.2 LE DOMAINE D'OLIGOMERISATION DES TET

---

Comme nous l'avons vu dans l'introduction de la thèse, le monomère de TET est composé d'un domaine catalytique et d'un domaine dit de dimérisation (Schoehn *et col.*, 2006). Le domaine de dimérisation a été identifié comme tel lors de la caractérisation de l'aminopeptidase PhTET1. En effet, lors de cette étude, après la caractérisation de la structure de PhTET1, une recherche par similarité structurale utilisant uniquement le domaine catalytique de l'aminopeptidase a retourné comme résultat la structure de l'AAP (pour *Aeromonas proteolyticus aminopeptidase*) (Schoehn *et col.*, 2006). Cette peptidase présente la particularité de posséder un domaine catalytique très proche de celui de PhTET1, mais montre une absence totale du domaine de dimérisation (figure 3-2). En effet, cette aminopeptidase de la famille M28 a été purifiée et caractérisée sous forme de monomère (Chevrier *et col.*, 1994).

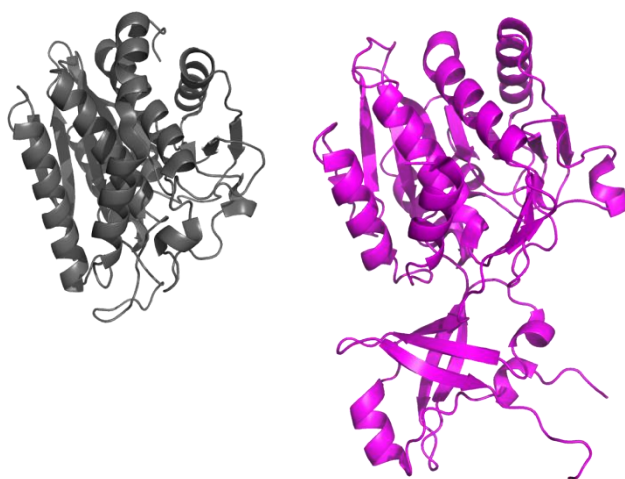


Figure 3-2 : Structure cristallographique d'AAP (code PDB : 1RTQ) en gris comparé au monomère de PhTET1 (code PDB : 2WYR) en magenta. Le domaine de dimérisation est absent chez l'aminopeptidase AAP.

Or, toutes les TET dodécamériques caractérisées possèdent le domaine supplémentaire de dimérisation. Sur la base de ces observations on peut proposer que la présence de ce domaine soit un marqueur de l'oligomérisation chez une peptidase de la famille M42. Cependant, comme nous allons le voir dans ce chapitre, l'oligomérisation de la particule TET ne peut être expliquée seulement par la présence de ce domaine. Il est donc nécessaire de déterminer des caractéristiques structurales plus précises afin de pouvoir identifier un assemblage de type TET.

### 3.3 LES INTERFACES D'OLIGOMERISATIONS

---

La dodécamérisation des aminopeptidases TET est permise par plusieurs interfaces d'oligomérisations : entre les monomères et entre les dimères. Des analyses structurales et l'utilisation de l'outil PISA (Sommaruga *et col.*, 2014) (voir la section 2.2) ont permis la caractérisation de ces zones d'interactions. L'outil d'analyse des interactions de PISA a permis de déterminer les résidus qui pourraient être impliqués dans les interactions entre monomères, dans la particule TET. Ces interactions potentielles sont le résultat de calculs de distances entre résidus à partir des structures cristallographiques et ne rendent donc pas compte des interactions en solutions. Cependant c'est un outil fiable pour proposer des interactions entre sous unités. L'objectif a été de confirmer ces hypothèses à l'aide de modèles structuraux. Pour cela, à l'aide de l'outil de recherche par homologie structurale Dali (Zaprasis *et col.*, 2014) (voir la section 2.2), des structures cristallographiques présentant des similarités structurales avec celle de PhTET2 (code PDB : 1Y0R) ont été recherchées dans la Protein Data Bank, avec un Z score minimum de 30.

Quatorze protéines dont les structures cristallographiques ont été résolues et présentant de fortes similarités structurales avec les TET déjà caractérisées ont été sélectionnées puis comparées en détails (Tableau 3-1). Plusieurs de ces protéines ont vu leurs structures résolues sans que des études biochimiques plus approfondies n'aient été réalisées. Néanmoins, en comparant les structures cristallographiques aux données obtenues sur leurs états d'oligomérisations lors de la purification de ces protéines, il est possible d'émettre des hypothèses quant à l'importance des différents domaines de la protéine pour le maintien de sa structure quaternaire.

L'ensemble des comparaisons structurales ne sera pas présenté et discuté dans ce chapitre. En effet, l'objectif est ici de définir les déterminants structuraux pour mieux caractériser une TET, de ce fait, les TET présentant une structure très proche de celle de PhTET1, PhTET2 ou PhTET3, et dont les sites de dimérisation et de dodécamérisation sont très similaires ne seront pas présentées en détail.

code PDB	Identité par rapport à PhTET2 (%)	Z-score	Organisme	Domaine	Type d'oligomérisation
1YOR	-	-	<i>Pyrococcus horikoshii</i>	Archée	dodécamère
1VHE	38	48,6	<i>Bacillus subtilis</i>	Bactérie	dodécamère
2WZN	50	48,5	<i>Pyrococcus horikoshii</i>	Archée	dodécamère
1YLO	30	45,3	<i>Shigella flexneri</i>	Bactérie	non publié
3KL9	34	45,3	<i>Streptococcus pneumoniae</i>	Bactérie	dodécamère
2WYR	37	43,6	<i>Pyrococcus horikoshii</i>	Archée	dodécamère
3ISX	36	37,9	<i>Thermotoga maritima</i>	Bactérie	non publié *
3CPX	26	36,9	<i>Cytophaga hutchinsonii</i>	Bactérie	trimère
2FVG	34	36,1	<i>Thermotoga maritima</i>	Bactérie	non publié
2GRE	22	34,8	<i>Bacillus cereus</i>	Bactérie	non publié
2GLF	15	33,7	<i>Thermotoga maritima</i>	Bactérie	non publié
4DYO	16	32,9	<i>Homo sapiens</i>	Eucaryote	dodécamère
4EME	14	32,8	<i>Plasmodium falciparum</i>	Eucaryote	dodécamère
4NJR	16	32,7	<i>Pseudomonas aeruginosa</i>	Bactérie	dodécamère
1VHO	31	32,5	<i>Thermotoga maritima</i>	Bactérie	monomère

Tableau 3-1 : Structures identifiées lors de la recherche d'homologues structuraux de l'aminopeptidase PhTET2 à l'aide de l'outil Dali. \* Présence uniquement du dimère dans l'unité asymétrique.

### 3.3.1 Interface de dimérisation

Cette zone est très conservée dans toutes les TET caractérisées. La disposition de ces interfaces et leurs natures suggèrent que le dimère, dont la formation implique un réseau fort d'interactions, pourrait être l'unité de base pour la formation du dodécamère. Cette question sera traitée dans le chapitre 4 de la thèse.

Cependant cette interface ne suffit pas pour identifier un édifice TET. On peut prendre pour exemple la protéine Q11Z05\_CYTH3 (uniprot) dont le code PDB est 3CPX (non publié).

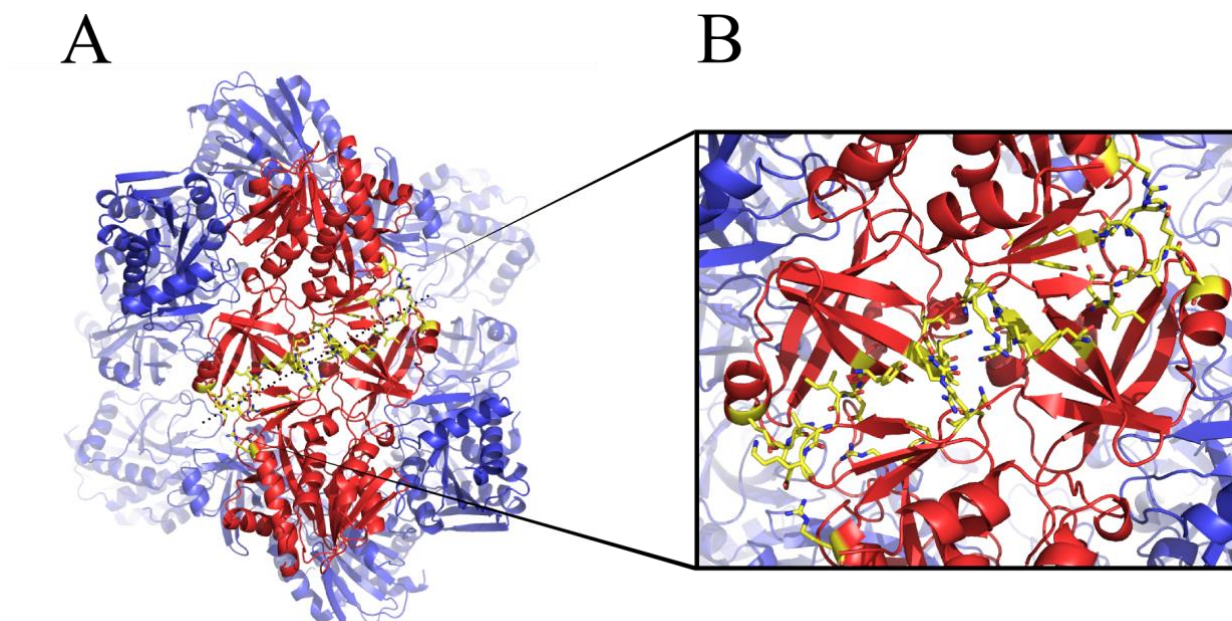


Figure 3-3 : Interface de dimérisation des TET. A) Représentation de la structure cristallographique de PhTET2 en ruban (Code PDB : 1Y0R). Un dimère formant une arête du tétraèdre est coloré en rouge. Les résidus impliqués dans des liaisons entre les monomères sont représentés en jaune. Une ligne pointillée représente l'axe de l'interface. B) Agrandissement de l'interface entre les monomères dans le dimère. Ce réseau d'interaction implique 17 résidus par monomère, on dénombre 28 liaisons hydrogènes et 20 ponts salins potentiels.

Cette protéine est assignée en tant qu'aminopeptidase frvX putative de la famille M42 et provient de la bactérie *Cytophaga hutchinsonii*. Elle présente une grande similarité structurale avec les TET et possède un domaine de dimérisation très semblable à celui des aminopeptidases TET de *P. horikoshii*.

Cependant, l'analyse de la structure cristallographique de cette protéine montre un trimère dans l'unité asymétrique. De plus, à partir de la symétrie du cristal, il n'est pas possible de retrouver une structure dodécamérique. Or, dans les cas de toutes les aminopeptidases TET dodécamériques caractérisées jusqu'à présent, il est possible d'obtenir la structure du dodécamère en appliquant les opérations de symétrie cristalline. Les données de purification ne sont pas publiées, cependant, sur la base de ces observations, on peut supposer que la protéine Q11Z05\_CYTH3 soit assemblée sous forme trimérique en solution.

L'analyse de la structure de Q11Z05\_CYTH3 révèle un assemblage trimérique formant une poche catalytique de TET (Figure 3-4 A et B). On remarque que la conformation du domaine de dimérisation est différente, de plus ce domaine est plus court que dans le

cas des autres TET, il est composé d'environ 60 résidus chez Q11Z05\_CYTH3 contre environ 80 résidus dans le cas des autres TET.

La particularité remarquable de la structure de Q11Z05\_CYTH3 est l'absence d'un motif « IDIGAXXE » situé à l'interface de dimérisation de TET. Ce motif est visible sur l'alignement des TET présenté en figure 3-11 entre les résidus 141 et 149. Ce motif est très conservé dans les autres TET analysées. Par ailleurs, l'arrangement général du domaine de dimérisation est légèrement différent entre PhTET2 et Q11Z05\_CYTH3 (Figure 3-4 C et E). Enfin, les résidus R320/Y321 dont les longues chaînes latérales entrent en jeu dans les contacts au sein du dimère ne sont pas conservés ici non plus (Figure 3-4D). Encore une fois ce motif est très conservé par ailleurs.

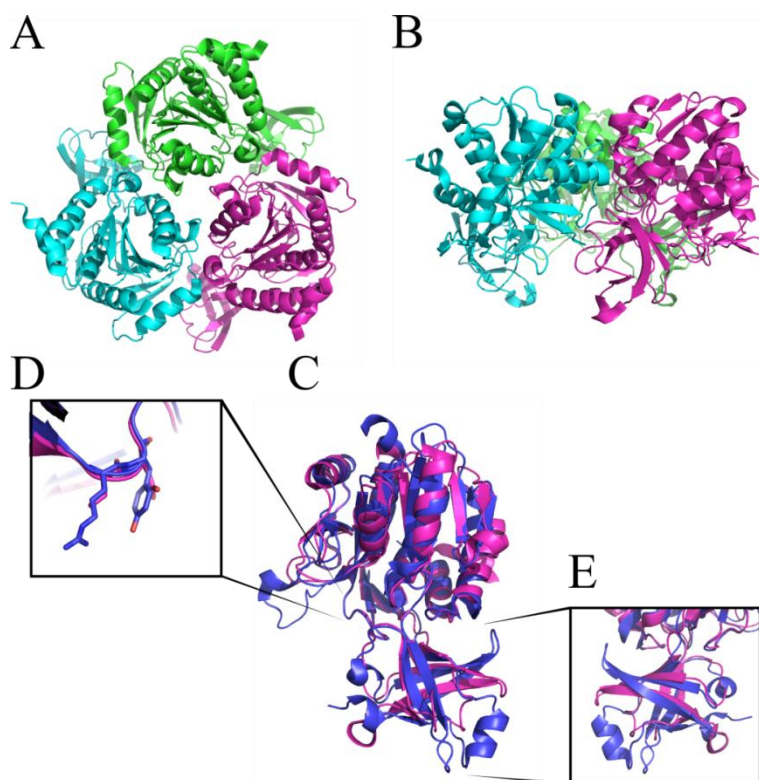


Figure 3-4 : Structure cristallographique de Q11Z05\_CYTH3. A) Structure cristallographique du trimère de Q11Z05\_CYTH3 présent dans l'unité asymétrique du cristal vu selon l'axe 3 de symétrie (code PDB : 3CPX). Chacun des monomères est représenté dans une couleur différente. B) Structure cristallographique du trimère de Q11Z05\_CYTH3 présent dans l'unité asymétrique du cristal vu selon un axe perpendiculaire à A). Chacun des monomères est représenté dans une couleur différente. C) Superposition des structures cristallographiques de PhTET2 (bleu foncé) (Code PDB : 1Y0R) et Q11Z05\_CYTH3 (rose). D) Agrandissement de la zone où sont présents les résidus R320/Y321 chez PhTET2. On remarque que les deux résidus sont mutés chez Q11Z05\_CYTH. E) Vue opposée à C) de la structure du domaine de dimérisation.

Cet exemple montre que la présence du domaine de dimérisation seul n'explique pas totalement l'oligomérisation de la particule. Dans le cas de Q11Z05\_CYTH3, l'ensemble de ces observations suggère que le motif « IDIGAXXE », associé à la présence des

résidus R(K)320/Y321, est essentiel pour la formation du dimère qui permet une dodécamérisation conduisant au complexe TET.

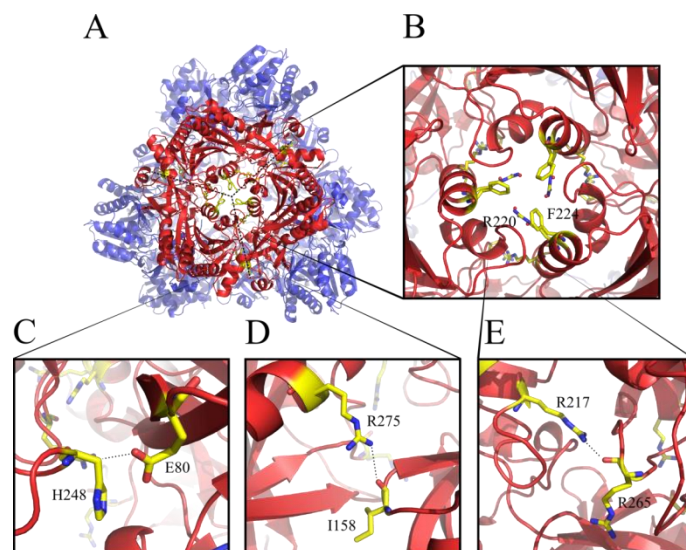
### 3.3.2 Interface de dodécamérisation

Les interactions entre les différents dimères, permettant la formation du dodécamère, sont principalement situées au niveau des « pointes », autrement dit les apices, du tétraèdre comme montré sur la figure 3-5A. Cette interface, dite de « dodécamérisation » est distincte de celle permettant la dimérisation. Elle est plus réduite et ne met en jeu qu'un nombre faible d'interactions. En première intuition, on aurait pu penser qu'un édifice tel que TET nécessite un grand nombre d'interactions pour être maintenue, il apparaît en réalité qu'il n'y en a que très peu (Figure 3-5). Cette zone de contact n'est pas très conservée au niveau des séquences. Mais on trouve des similarités structurales fortes. Les résidus placés le long de l'hélice  $\alpha 5$  formé par les résidus 216 à 227 (selon la nomenclature des structures secondaires de PhTET2, code PDB : 1Y0R, voir alignement figure 3-11) à l'interface entre les sous-unités formant la poche catalytique contiennent toujours une séquence de 3 résidus à longue chaîne latérale chargée (positivement ou négativement), ces dernières étant orientées vers une autre sous-unité de l'interface (Figure 3-5B). Je pense donc que cette zone est importante à la fois pour la structure de la TET, mais aussi pour son activité. La nature des résidus présents pourrait avoir une incidence sur la nature de l'activité de la TET. En effet, les apices de la protéine TET représentent des « trous de sortie » pour les résidus hydrolysés, la nature des acides aminés qui les composent pourrait donc avoir une incidence sur l'efficacité avec laquelle les résidus hydrolysés sont évacués de la poche catalytique. Cette possibilité reste encore à explorer.

L'importance des résidus de l'hélice  $\alpha 5$  pour la formation de l'assemblage TET est confortée par l'analyse de la structure de la protéine TmPep1050 provenant de la bactérie hyperthermophile *Thermotoga maritima* dont le code PDB est 3ISX (non publié). Elle présente 38 % de similarité de séquence avec PhTET2 (Dutoit *et col.*, 2012). L'outil de comparaison de structure DALI calcule un Z score de 38.2 et les deux structures se superposent bien (Figure 3-6A). La particularité de cette protéine est qu'elle a été purifiée sous forme de dimère exclusivement (Craik *et col.*, 2011). Dans le



cas de TmPep1050, il n'est pas possible non plus de reconstruire un dodécamère en utilisant la symétrie cristalline.



**Figure 3-5 : Interface de dodécamérisation des TET.** A) Représentation de la structure cristallographique de PhTET2 en ruban (Code PDB : 1Y0R). Un trimère formant un apex du tétraèdre est coloré en rouge. Les résidus impliqués dans des liaisons faibles entre les monomères sont représentés en jaune. Une ligne pointillée représente les axes des interfaces. Ce réseau, moins bien défini que l'interface dimérique, implique au moins 5 résidus par monomère, on dénombre 2 liaisons hydrogènes et 1 pont salin potentiel. B) Agrandissement du pore situé à la pointe du tétraèdre, les résidus R220 et F224 sont impliqués dans un réseau d'interaction polaire et hydrophobe potentiel. C) Agrandissement de l'interaction de type pont salin potentiel entre les résidus H248 et E80 détecté par DALI. D, E) Agrandissement des interactions de type liaison hydrogène potentiel entre les résidus R275 et R265 d'un monomère et I158 et R217, respectivement, d'un second monomère.

L'analyse de la structure montre que l'interface supposée de dodécamérisation est ici fortement modifiée (Figure 3-6 B et C). On remarque que l'hélice  $\alpha 5$  portant les résidus impliqués dans les contacts entre les monomères aux apex dans les TET caractérisées n'est ici plus du tout structurée (Figure 3-6C). Les résidus à longue chaîne latérale pointant vers l'interface ne sont pas présents chez TmPep1050, on observe les résidus valine, sérine et tyrosine, orientés vers l'intérieur de la particule, à la place des trois résidus retrouvés dans les autres structures de TET (Figure 3-6C et alignement en Figure 3-11).

Par ailleurs, on peut aussi remarquer que l'hélice alpha adjacente à l'interface d'oligomérisation, représentée en Figure 3-6B, n'est pas orientée de la même façon que dans PhTET2. En effet, dans le cas de TmPep1050, celle-ci est décalée vers l'extérieur de la particule et empêche de toute évidence la formation de dodécamère.



L'ensemble de ces observations suggère que l'hélice  $\alpha 5$  située à l'interface entre les monomères aux apex du dodécamère, ainsi que la présence des résidus à longue chaîne latérale pointant vers l'interface sont essentiels pour la formation du dodécamère.

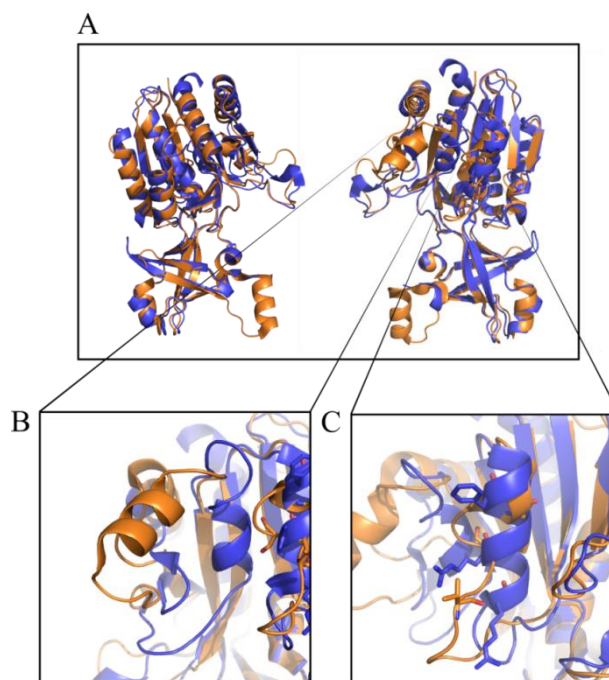


Figure 3-6 : Comparaison des structures cristallographiques de PhTET2 et TmPep1050. A) Vues opposées de la superposition des structures cristallographiques de PhTET2 (bleu foncé) (Code PDB : 1Y0R) et TmPep1050 (orange) (Code PDB : 3ISX). B) Agrandissement d'une hélice alpha adjacente à l'interface entre les monomères à l'apex du dodécamère. On remarque la modification de l'orientation de cette hélice. C) Agrandissement de l'hélice alpha  $\alpha 5$  situé à l'interface entre les monomères à l'apex du dodécamère. On remarque la modification de l'arrangement de celle-ci chez TmPep1050 et l'absence de résidu à longue chaîne carbonée.

#### 3.3.3 Cas particulier de Tm1048

Le dernier exemple de structure présentant des similarités structurales avec PhTET2 est la peptidase putative Tm1048 isolée chez *Thermotoga maritima* dont le code PDB est 1VHO (Krupovic *et col.*, 2014). Cette peptidase de la famille M42 dont la structure est résolue est présente dans l'unité asymétrique du cristal sous forme de monomère.

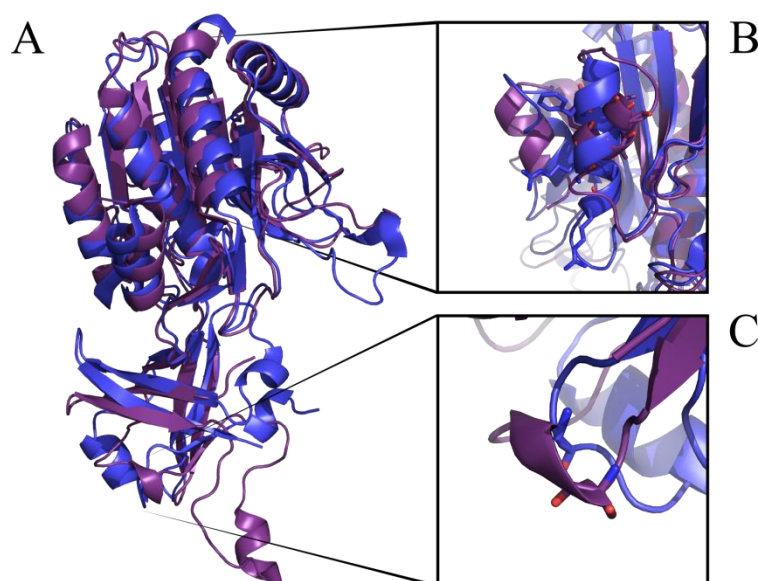


Figure 3-7 : Superposition des structures cristallographiques de PhTET2 (Code PDB : 1Y0R) et Tm1048 (Code PDB : 1VHO). A) Superposition des structures cristallographiques des monomères de PhTET2 (bleu foncé) et Tm1048 (mauve). B) Agrandissement de l'hélice  $\alpha 5$  à l'interface entre les monomères à l'apex du dodécamère. On remarque la modification de l'arrangement de celle-ci chez Tm1048 et l'absence de résidu à longue chaîne latérale. C) Agrandissement de la zone du motif « IDIGAXXXE ». La glycine est ici remplacée par une sérine, modifiant ainsi l'arrangement de cette zone importante pour la dimérisation.

Sa structure, encore une fois très proche de celle des TET de *P. horikoshii* présente toutefois quelques différences (Figure 3-7A). L'hélice  $\alpha 5$  formant l'interface entre les monomères aux apices du dodécamère de TET est ici fortement perturbée, de plus, aucun résidu à longue chaîne latérale n'est présent (Figure 3-7B). On remarque également une modification du domaine « IDIGAXXXE » que l'on vient d'identifier comme un motif essentiel pour la dimérisation des TET (Figure 3-7C). En effet, la glycine de ce motif est ici remplacée par une sérine qui en modifie la structure. On retrouve cependant les résidus R320/Y321 jouant également un rôle dans l'interface entre les dimères.

La visualisation rapide de la structure du monomère de Tm1048 suggérerait que c'est une sous unité d'un complexe TET, pourtant des modifications fines de la structure rendent impossible l'assemblage en dodécamère de cette protéine. Cet exemple rend compte de l'importance de la spécification des déterminants structuraux pour l'identification des TET.

### 3.4 DETERMINANTS STRUCTURAUX SPECIFIQUES DE LA FONCTION DES PEPTIDASES TET

#### 3.4.1 Les glycines : pivots de la structure

L'analyse des séquences des protéines dont les structures présentent une forte similarité avec les TET de *P. horikoshii* révèle la présence de plusieurs résidus glycine très conservés.

Ces résidus glycine sont situés entre différents feuillets beta et hélices alpha. Ils apportent la flexibilité nécessaire au repliement de la structure (Figure 3-8). Les glycines 23, 215 et 218 confèrent une flexibilité nécessaire au repliement du domaine catalytique à proximité du site actif (numérotation des résidus PhTET2). Les glycines 73, 91 et 92 sont, quant à elles, situées à la jonction entre le domaine catalytique et le domaine de dimérisation. L'ensemble des glycines 23, 73, 91, 92, 215 et 218 sont situées à proximité du site catalytique. Ces observations suggèrent que ces résidus glycines conservés sont importants pour le repliement de la structure autour du site actif et donc pour l'activité de l'aminopeptidase.

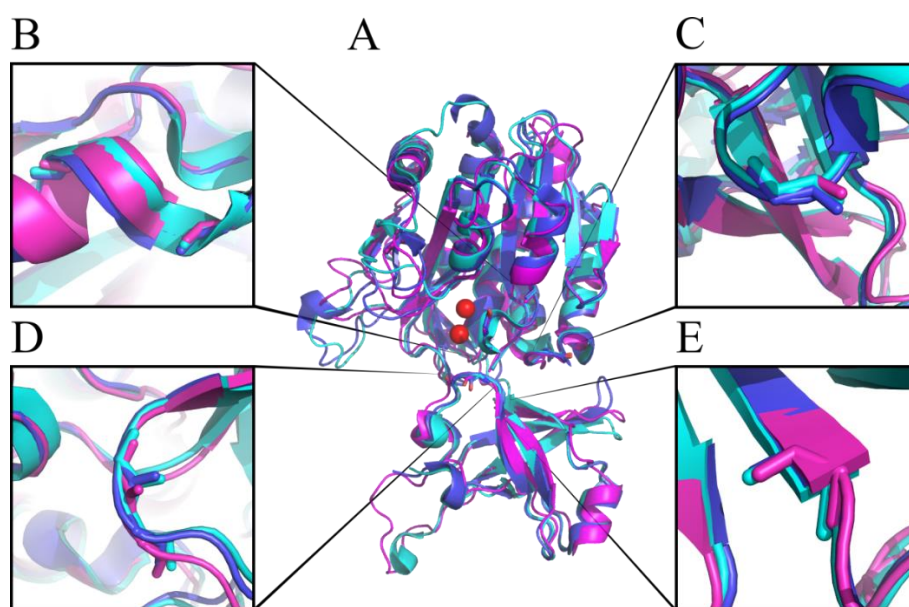


Figure 3-8 : Superposition des monomères de PhTET et agrandissement des zones où se trouvent les glycines conservées. A) Superposition des monomères de PhTET1 (magenta) (Code PDB : 2WYR), PhTET2 (bleu foncé) (Code PDB : 1Y0R) et PhTET3 (cyan) (Code PDB : 2WZN), les ions métalliques du site actif sont représenté par les sphères rouge B) Glycine 23. C) Glycine 215XX218. D) Glycine 91-92. E) Glycine 73. (Numérotation des résidus de PhTET2).

#### 3.4.2 Conservation du site actif

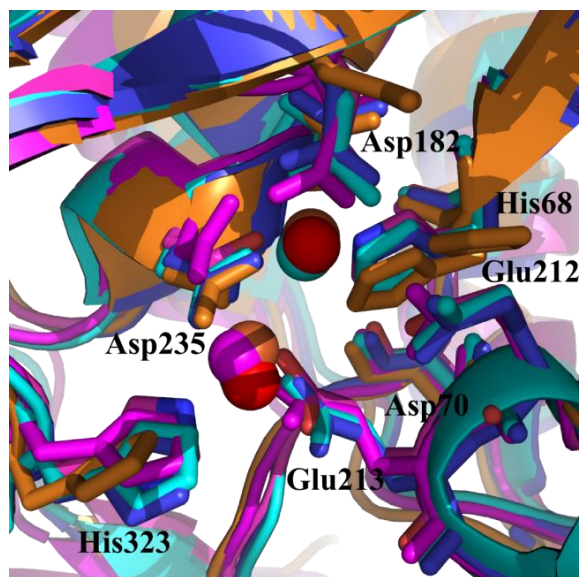


Figure 3-9 : Superposition des structures conservées du site actif de plusieurs TET. (PhTET1 : magenta ; PhTET2 : bleu foncé ; PhTET3 : cyan ; PepA : orange) (Numérotation des résidus PhTET2)

Comme décrit dans l'introduction (voir section 1.2), la classification des peptidases est basée sur la conservation des sites actifs associés à l'activité de la famille de la peptidase. Dans l'étude décrite ici, plusieurs protéines dont la structure est très similaire à celle des aminopeptidases TET ont été analysées.

Chacune de ces protéines présente un site actif très conservé et caractéristique des métallopeptidases de la famille M42

### 3.5 SYNTHÈSE : LES MOTIFS STRUCTURAUX POUR LA DETERMINATION DES COMPLEXES TET

---

L'ensemble des observations faites jusqu'à présent permet de définir plusieurs motifs structuraux conservés chez les TETs pour l'identification d'une peptidase comme appartenant à cette famille de machines macromoléculaires (Figure 3-10). Nous avons retenu 5 critères:

- 1) **L'insertion du domaine de dimérisation dans la séquence** : cette insertion d'un domaine spécifique au milieu du gène codant pour l'aminopeptidase est un premier marqueur fort de la structure des aminopeptidases TET (Schoehn *et col.*, 2006).
- 2) **L'interface de dimérisation** : La présence du motif « IDIGAXXE » paraît être déterminante pour l'assemblage des dimères de TET. La présence des résidus R320Y321 paraît être elle aussi importante (Figure 3-10C, D et E).

3) **L'interface de dodécamérisation** : c'est une zone où les interactions sont moins bien définies, elle implique des interactions polaires, hydrophobes, des liaisons hydrogènes et un pont salin. La présence de l'hélice  $\alpha 5$  à l'interface entre les monomères dans une conformation adéquate associée à la présence de résidus à longue chaîne latérale chargée sur cette dernière semble être déterminante. Il apparaît cependant difficile de déterminer une séquence type qui pourrait être identifiée dans les génomes (Figure 3-10B).

4) **Le repliement du domaine catalytique** : la présence de plusieurs glycines très conservées semble être importante pour le repliement du domaine catalytique des TET (Figure 3-8).

5) **Le site actif** : La conservation du site fait partie des caractéristiques d'une TET et des peptidases de la famille M42, il est défini par les résidus suivant : H65XD67, D181, E213E214, E/D236 et H319 chez PhTET3 et implique la coordination de deux ions métalliques. Ces ions sont, en général, des ions  $\text{Co}^{2+}$  ou  $\text{Zn}^{2+}$ , mais certaines structures ont été caractérisées avec des ions différents, la structure de Q11Z05\_CYTH3 a par exemple été caractérisée avec des sites actifs portant des ions  $\text{Fe}^{2+}$ .

Ces cinq critères doivent être réunis pour identifier une peptidase en tant que peptidase tétraédrique TET.



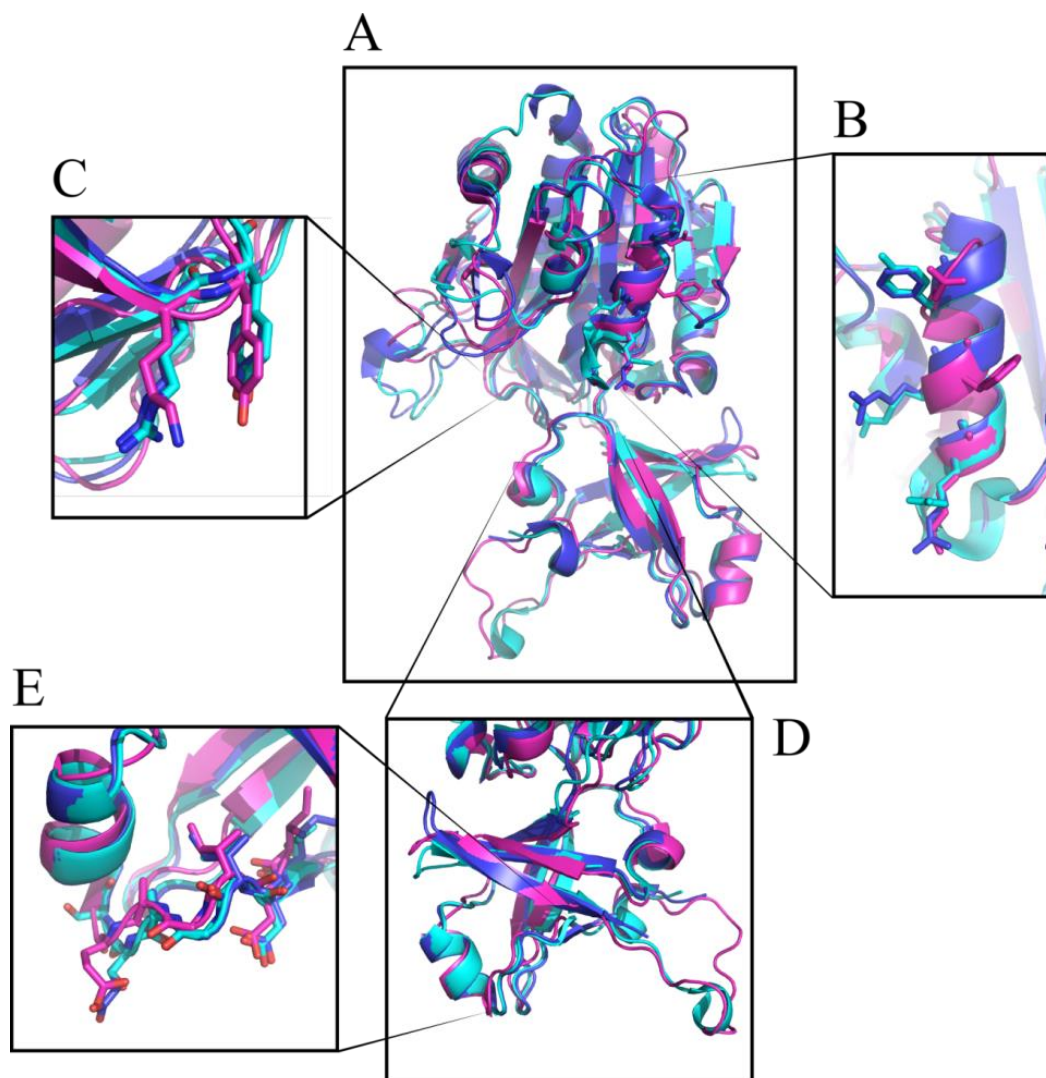


Figure 3-10 : Interfaces d'oligomérisation des TET : A) Superposition des structures cristallographiques des monomères de PhTET1 (magenta) (Code PDB : 2WYR), PhTET2 (bleu foncé) (Code PDB : 1Y0R) et PhTET3 (cyan) (Code PDB : 2WZN). B) Agrandissement de l'hélice  $\alpha 5$  impliquée dans l'interface de dodécamérisation. Les acides aminés à longue chaîne latérale conservés dans cette zone sont représentés en bâtonnets. C) Agrandissement des résidus R352Y353 (PhTET2 ou PhTET3) ou K352Y353 (PhTET1) impliqués dans l'interface de dimérisation. D) Agrandissement du domaine de dimérisation en vue opposée. E) Agrandissement du motif IDIGAXXE impliqué dans l'interface de dimérisation. Les résidus du motif sont représentés en bâtonnets.

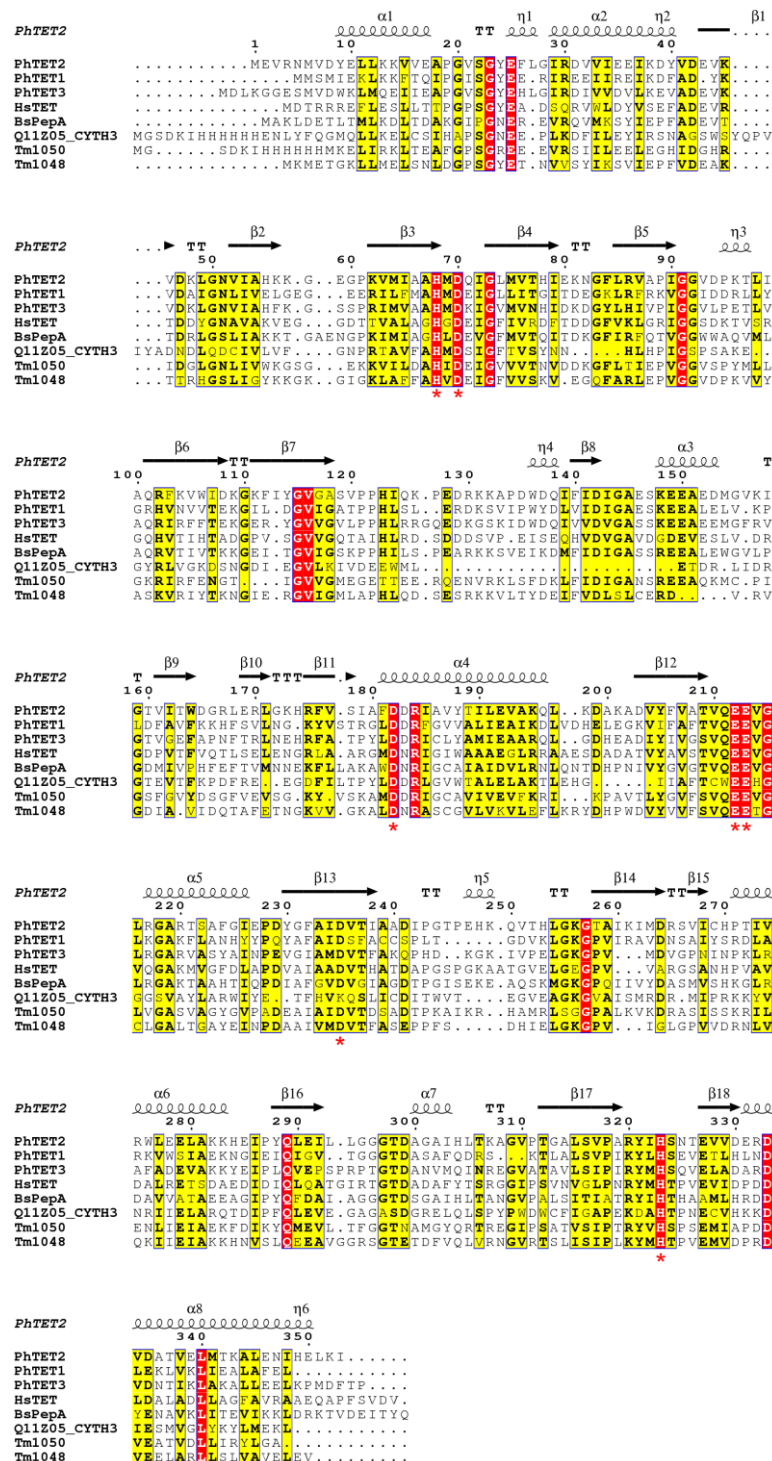


Figure 3-11 : Alignement des séquences d'aminopeptidases TET. La numérotation des résidus correspond à la séquence de PhTET2. Les résidus du site actif sont indiqués par une étoile rouge. Dans le cas de l'Asp 235, il apparaît sur l'alignement que la peptidase Q11Z05\_CYTH3 ne le possède pas, en réalité, la superposition des structures montre que cet Asp est bien présent dans le site actif et correspond à l'Asp en position 241 sur la séquence de Q11Z05\_CYTH3, son site actif est donc bien conservé. PhTET2, PhTET1, PhTET3 : TET2, TET1 et TET3 de *Pyrococcus horikoshii* ; HsTET : TET de *Halobium salinarum* ; BsPepA : PepA de *Bacillus subtilis* (TET bactérienne) ; Q11Z05\_CYTH3 : TET potentielle de *Cytophaga hutchinsonii* ; Tm1050 et Tm1048 : TET potentielle de *Thermotoga maritima*. La numérotation des résidus correspond à PhTET2, les structures secondaires indiquées sont celles de PhTET2

## 3.6 PHYLOGENIE DES TET CHEZ LES ARCHEES



Les déterminants structuraux que nous avons identifiés vont permettre de réaliser en collaboration avec Simonetta Gribaldo à l'Institut Pasteur, à Paris, une étude phylogénétique sur les peptidases TET dans les 3 domaines du vivant.

Pendant la thèse, une étude préliminaire a été faite par S. Gribaldo chez les Archées. Les séquences de TET déjà caractérisées, PhTET1, PhTET2 et PhTET3, ont été utilisées pour identifier des homologues de séquences dans l'ensemble des génomes Archées. Nous avons ainsi obtenu une phylogénie des TET chez les Archées dont une représentation schématique est présentée en figure 3-12. Le travail va être affiné avec les données issues des analyses structurales présentées dans ce chapitre. Cette étude a cependant permis de préciser les orientations scientifiques de cette thèse.

Plusieurs résultats très intéressants ressortent de cette étude.

Premièrement, les TET sont effectivement très conservées dans les génomes d'Archées. Il apparaît cependant qu'en fonction des ordres d'Archées, le nombre de TET varie de zéro à quatre TET. Les *Halobacteriales* peuvent en posséder jusqu'à 3. Certains ordres en possèdent deux : les *Nanohaloarchaea*, la *Korarchaeota* ainsi que les *Thermoproteales*. La majorité des ordres en possèdent une seule et certains n'en ont aucune (Figure 3-12). Enfin, l'ordre des *Thermococcales* est le seul à posséder jusqu'à 4 TET. Cette étude a donc permis de proposer l'existence d'une quatrième TET chez *P. horikoshii*. J'ai commencé sa caractérisation durant ma thèse, les résultats préliminaires obtenus sont présentés dans le chapitre 7 de ce manuscrit.

Par ailleurs, cette étude a également révélé que les TET des *Thermococcales* sont en réalité issues de plusieurs transferts de gènes horizontaux depuis les Bactéries. Les TET des *Halobacteriales*, de la *Korarchaeota* ainsi que celle du groupe II en sont également issues.

Cette étude montre aussi l'existence d'une TET d'origine « Archéenne » qui devait être présente dans l'ancêtre commun des Archées (LACA).

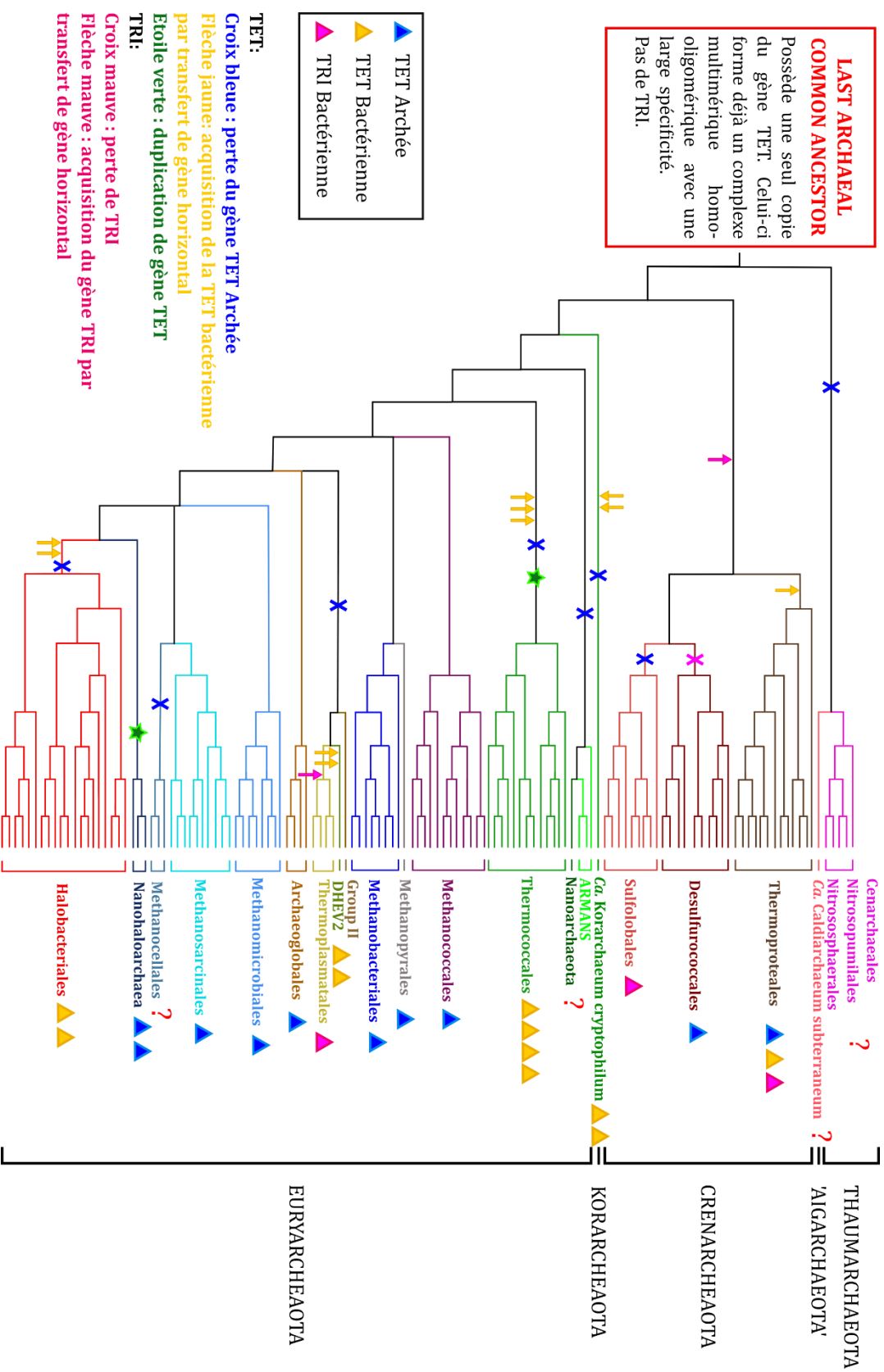


Figure 3-12 : Représentation schématique de la phylogénie des TET et TRI chez les Archées. La description des codes couleur est indiquée sur la figure. La phylogénie schématisée est basée sur une analyse bayésienne de protéine ribosomale concaténée.

Celle-ci est généralement présente en une seule copie, sauf chez les *Nanohaloarchaea*, qui en possèdent deux. Les ordres ne possédant aucune TET ont donc « perdu » cette TET originale. Dans le laboratoire ELMA, l'organisme modèle utilisé pour l'étude de cette TET est l'Archaea *Methanocaldococcus janashii*.

Enfin ce travail a permis de mettre en évidence l'exclusion mutuelle des systèmes TET et TRI (Tricorn) dans les génomes d'Archées. Nous rappelons que l'activité de TRI (voir section 1.3.1), couplé à ses cofacteurs F1, F2 et F3, a une action proche de celle des TET, à savoir l'hydrolyse complète de peptides en acides aminés libres. Il est donc intéressant de noter que les ordres des *Thermoplasmatales* et des *Sulfolobales*, qui possèdent TRI, ne possède pas de TET. A l'inverse les ordres possédant une TET n'ont pas de TRI. La seule exception étant l'ordre des *Thermoproteales* où, étonnamment, une TET bactérienne, une TET archéenne et la peptidase TRI co-existent dans la même cellule.

### 3.7 DISCUSSION

---

Le travail d'analyse structurale présenté dans ce chapitre associé aux études phylogénomiques renforce l'intérêt des travaux de biochimie et de biophysique présentés dans les chapitres suivants. Tout d'abord l'analyse souligne que les TETs sont des machines cellulaires très répandues chez les procaryotes. Ceci indique que leur rôle physiologique fondamental n'est pas restreint à une fonction "marginale", spécifique à une classe de microorganisme. Il apparaît aussi que la multiplicité des TET n'est pas une constante et est une caractéristique très marquée chez les Archées hyperthermophiles de l'ordre des *Thermococcales*. L'étude sur les propriétés d'hétéro-oligomérisation présentée dans le chapitre 5, ainsi que la caractérisation préliminaire du complexe TET4 présenté dans le chapitre 7 donne une explication partielle à cette particularité. L'étude comparative des structures de peptidases M42 montre également que l'oligomérisation est une caractéristique qui n'est pas "portée" par le domaine catalytique, mais semble avoir été acquise via l'insertion d'un domaine spécifique ainsi que par des motifs spécifiques de la dodécamérisation. Ceci renforce l'intérêt des études présentées dans les chapitres 5 et 6 sur les processus d'assemblages et leurs conséquences sur la biologie du système TET.



## 4. LES AMINOPEPTIDASES TET2 ET TET3 DE *PYROCOCCUS HORIKOSHII* FORMENT UN PEPTIDASOME A SOUS-UNITES MULTIPLES DOTE DE PROPRIETES DE DESTRUCTION DES POLYPEPTIDES AMELIOREES.

### 4.1 PRESENTATION DU TRAVAIL

---

L'analyse de la présence de peptidases TET dans les cellules d'Archées présentée dans le chapitre précédent a montré qu'elles peuvent avoir des origines évolutives diverses et provenir de multiples évènements de transfert de gènes, en particulier à partir des bactéries. Les raisons de cette multiplicité de peptidases TET, en particulier chez les *Thermococcales*, sont mal connues. Dans le cas des 3 TETs de *Pyrococcus horikoshii* (nombre qui sera probablement porté à 4 une fois la caractérisation de *P.horikoshii* achevée) l'hypothèse formulée dans l'article de Durá *et col.* (2009) était que les 3 machines peptidasiques PhTET1, PhTET2 et PhTET3 représentent un système intégré de protéolyse, issu d'un ancêtre commun, en conjuguant leurs spécificités de clivage des acides aminés N-terminaux. L'analyse de S. Gribaldo de l'Institut Pasteur indiquait plutôt des évènements multiples et une origine bactérienne ainsi qu'une origine commune pour TET2 et TET3. Dans ces conditions, la question de la capacité de ces différentes peptidases à former des complexes hétéro-oligomériques se posait avec plus d'insistance.

Dans l'article présenté dans ce chapitre, j'ai utilisé la méthode du fractionnement cellulaire par ultracentrifugation de gradients de sucrose. Cette technique permet, avec une bonne calibration des fractions réalisée avec des protéines de poids moléculaires connus, d'apprécier l'état d'oligomérisation d'une protéine dans un contexte cellulaire. Elle a aussi été utilisée dans l'article présenté dans le chapitre 5 pour montrer que les formes de bas poids moléculaires qui s'accumulent dans les Archées halophiles et dans *Pyrococcus horikoshii* sont des dimères. Pour ces expériences, nous avons dû produire des anticorps spécifiques de chacune des aminopeptidases TET. Ces anticorps sont

dirigés contre des peptides non conservés entre les 3 TETs. J'ai vérifié qu'ils ne présentaient pas de réactions croisées entre eux ni avec d'autres protéines de *Pyrococcus horikoshii*.

Nous avons ensuite étudié les protéines TET obtenues après co-expression dans *E. coli* grâce au plasmide PDuet. Ce travail est basé sur les observations d'A. Durà, post-docorante au laboratoire qui avait réalisé les études pilotes sur l'hétéro-oligomérisation des TET et avait montré qu'il était possible d'utiliser des colonnes échangeuses d'ions à haute résolution pour séparer les différents types d'assemblages TET en fonction de leurs charges de surface. Seule la co-expression PhTET2-PhTET3 permet de purifier des formes présentant un volume d'élution différent de celui des homo-oligomères PhTET2 ou PhTET3. En effet, les co-expressions de PhTET1 avec les autres PhTET conduisent systématiquement à la formation d'agrégats. Les analyses en filtration sur gel et en microscopie électronique ont montré qu'il s'agissait bien de complexes TET dodécamériques. Ces complexes combinent les activités spécifiques de PhTET2 et PhTET3. Ceci indiquait donc que nous étions en présence de complexes hétéro-oligomériques.

Nous avons réussi à purifier à homogénéité différentes espèces de complexes TET résultant de la co-expression PhTET2-PhTET3. Celles-ci présentent exactement le même comportement chromatographique que les espèces détectées par immunodétection dans les extraits de *Pyrococcus horikoshii*, montrant ainsi l'existence de complexes hétéro-oligomériques dans les cellules de *Pyrococcus horikoshii*.

Pour comparer l'efficacité enzymatique des complexes hétéro-oligomériques vis à vis des formes homo-oligomériques, nous devons déterminer la stœchiométrie de ces assemblages. Nous avons tout d'abord tenté de la déterminer par spectrométrie de masse. Malheureusement les différences de masse entre les monomères PhTET2 et PhTET3 étaient trop faibles pour obtenir des résultats fiables. J'ai ensuite tenté de quantifier les proportions de sous-unités PhTET2 par rapport aux sous-unités PhTET3 en utilisant la méthode du Western blot quantitatif qui consiste à transférer sur une même membrane l'échantillon avec une gamme de la protéine d'intérêt. Là encore, les résultats n'étaient pas assez précis. Finalement, nous avons utilisé les complexes hétéro-oligomériques obtenus par reconstruction *in vitro* pour déterminer la composition des

complexes obtenus *in vivo*. Ces expériences de reconstruction décrites dans le chapitre 5 ont permis de déterminer la topologie précise des sous-unités au sein des complexes PhTET2:PhTET3 par les études de diffusion neutronique aux petits angles. Ces études nous ont permis de caractériser des complexes iso-stœchiométriques de PhTET2:PhTET3, c'est-à-dire comportant 6 sous-unités de PhTET2 et 6 sous-unités de PhTET3. En chromatographie échangeuse d'ions à haute résolution, ces complexes se comportent précisément comme la forme majoritaire qui s'accumule *in vivo*, dans *E. coli* après co-expression, ou dans les cellules de *Pyrococcus horikoshii*.

Nous avons ensuite étudié l'efficacité enzymatique du complexe hétéro-oligomérique iso-stœchiométrique. Pour cela nous avons eu l'idée d'utiliser des peptides chromogéniques comportant une séquence d'acides aminés coupés alternativement uniquement par PhTET2 ou par PhTET3. Ainsi les complexes homo-oligomériques seuls n'ont pas d'activité observable sur ces substrats puisque il faut une destruction totale du peptide pour libérer le chromophore situé en C-ter. Plusieurs substrats ont été ainsi utilisés et ont donné des résultats concordants. Seule une de ces expériences est présentée dans l'article. Nous avons également purifié les acides aminés produits par ces réactions et mesuré leur cinétique de libération en les analysant sur une colonne d'affinité. Enfin, j'ai aussi étudié par HPLC en phase inverse sur une colonne d'interaction hydrophobe  $\mu$ RPC de type C2/C18 l'efficacité de dégradation des différents types ou mélange de peptidases PhTET, homo-oligomériques ou hétéro-oligomériques sur des mélanges de peptides complexes.

Cet article comprend également une analyse des interfaces entre les peptidases TET permettant d'expliquer l'absence d'interaction de PhTET2 et PhTET3 avec PhTET1 ainsi qu'une étude phylogénomique sur les 3 protéines TETs des *Thermococcales*, réalisée par S. Gribaldo qui permet de proposer un scénario évolutif expliquant la relation fonctionnelle entre TET2 et TET3.



**4.2 "THE TET2 AND TET3 AMINOPEPTIDASES FROM PYROCOCCLUS  
HORIKOSHII FORM A HETERO-SUBUNIT PEPTIDASOME WITH  
ENHANCED PEPTIDE DESTRUCTION PROPERTIES".**

---

# The TET2 and TET3 aminopeptidases from *Pyrococcus horikoshii* form a hetero-subunit peptidasome with enhanced peptide destruction properties

Alexandre Appolaire,<sup>1,2,3</sup> M. Asunción Durá,<sup>1,2,3</sup>  
Mylène Ferruit,<sup>1,2,3</sup> Jean-Pierre Andrieu,<sup>1,2,3</sup>  
Anne Godfroy,<sup>4</sup> Simonetta Gribaldo<sup>5</sup> and  
Bruno Franzetti<sup>1,2,3\*</sup>

<sup>1</sup>Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France.

<sup>2</sup>CNRS, IBS, F-38027 Grenoble, France.

<sup>3</sup>CEA, DSV, IBS, F-38027 Grenoble, France.

<sup>4</sup>Ifremer, UMR6197, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France.

<sup>5</sup>Institut Pasteur, Unité Biologie Moléculaire du Gène chez les Extremophiles, Département de Microbiologie, 75724 Paris Cedex 15, France.

## Summary

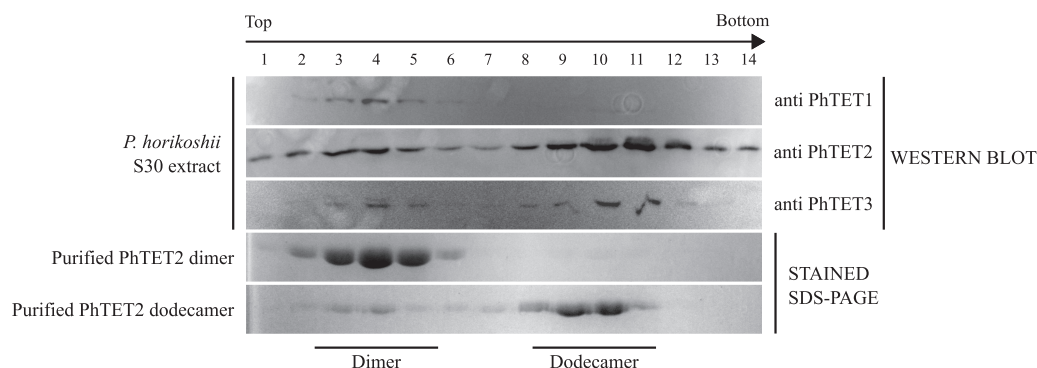
**TET aminopeptidases assemble as large homo-dodecameric complexes. The reason why prokaryotic genomes often encode a diverse set of TET peptidases homologues remains unclear. In the archaeon *Pyrococcus horikoshii*, PhTET1, PhTET2 and PhTET3 homo-oligomeric particles have been proposed to work in concert to breakdown intracellular polypeptides. When coexpressed in *Escherichia coli*, the PhTET2 and PhTET3 proteins were found to assemble efficiently as heteromeric complexes. Biophysical analysis demonstrated that these particles possess the same quaternary structure as the homomeric TET dodecamers. The same hetero-oligomeric complexes were immunodetected in *P. horikoshii* cell extracts analysed by sucrose gradient fractionation and ion exchange chromatography. The biochemical activity of a purified hetero-oligomeric TET particle, assessed on chromogenic substrates and on a complex mixture of peptides, reveals that it displays higher efficiency than an equivalent combination of homo-oligomeric TET particles. Interestingly, phylogenetic analysis shows that PhTET2 and PhTET3 are paralogous proteins that arose from gene duplication in the ancestor**

**of Thermococcales. Together, these results establish that the PhTET2 and PhTET3 proteins are two subunits of the same enzymatic complex aimed at the destruction of polypeptidic chains of very different composition. This is the first report for such a mechanism intended to improve multi-enzymatic complex efficiency among exopeptidases.**

## Introduction

In microorganisms, intracellular peptidases are important for nutrition, stress response and pathogenicity (Gonzales and Robert-Baudouy, 1996; Ingmer and Brondsted, 2009). TET peptidases are large cytosolic complexes that possess a unique tetrahedral quaternary structure (Franzetti *et al.*, 2002; Durá *et al.*, 2009). They are composed of 12 subunits and have the active sites located in the interior of the molecule. The X-ray structures of different homo-oligomeric TET particles, combined with electron microscopy studies, have revealed an elaborated self-compartmentalized organization (Russo and Baumann, 2004; Schoehn *et al.*, 2006; Durá *et al.*, 2009). The access of substrate polypeptides is restricted by four pores situated on each facet of the tetrahedron. The interior of the complex contains four channels leading to the apices, where each of the four proteolytic chambers contains three active sites. TET particles were first described in the domain of Archaea (Franzetti *et al.*, 2002), but TET homologues are also widely present in bacterial genomes, including pathogens such as *Streptococcus pneumonia* (Lutfullah *et al.*, 2009; Kim *et al.*, 2010). TET peptidases belong to the M42 family of the clan MH of metallopeptidases in the MEROPS classification system (Rawlings *et al.*, 2012). Their self-oligomerization is due to an additional dimerization domain that is not present in monomeric or secreted aminopeptidases from the same clan (Schoehn *et al.*, 2006). Since in many peptidases, active-site internalization is required to prevent unwanted proteolytic damage against cytosolic polypeptides, it was inferred that the dimerization domain was added during evolution to allow cytosolic localization of the TET peptidases. In fact, we recently showed that dimers represent

Accepted 25 August, 2014. \*For correspondence. E-mail franzetti@ibs.fr; Tel. (+33) 4 38 78 95 69; Fax (+33) 438 78 54 94.



**Fig. 1.** Determination of TET oligomerization states in *P. horikoshii* cellular extracts. Rows 1–3: Immunodetection of native PhTET1, PhTET2 and PhTET3 complexes from *P. horikoshii* cellular extracts. Post-membrane extracts (S30) of *P. horikoshii* were loaded on a 5–25% sucrose density gradient and resolved by ultracentrifugation. The top arrow indicates the sedimentation direction. The fractions numbers are indicated on the top. The proteins of the different fractions were loaded on 12% SDS-PAGE gels and immunodetected using specific antibodies. Rows 4–5: Purified recombinant dimer and dodecamer of PhTET2 separated by 5–25% sucrose density gradients. The fractions were loaded on 12% SDS-PAGE gels and revealed with Coomassie blue staining. Dimer and dodecamer of PhTET2 and PhTET3 are accumulated *in vivo*, while PhTET1 is only present under its dimeric form.

the precursors in the TET assembling pathway and that self-oligomerization triggers the enzymatic efficiency of the system (Appolaire *et al.*, 2013). Interestingly, stable dimeric precursors have been detected in the cytosol together with assembled dodecamers, suggesting the existence of a functional regulation based on oligomerization control *in vivo* (Appolaire *et al.*, 2013).

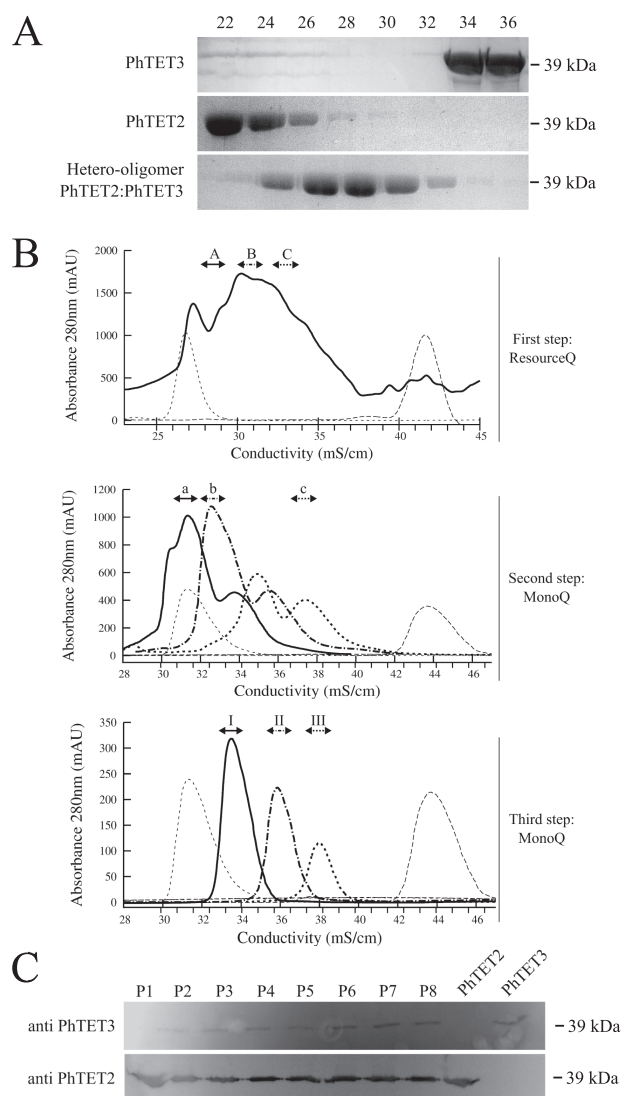
It is not uncommon that archaeal and bacterial genomes encode for several TET homologues. The structures and functions of three TET dodecameric complexes from the deep-sea hyperthermophilic archaeon *Pyrococcus horikoshii* (PhTET1, PhTET2 and PhTET3) have been determined (Russo and Baumann, 2004; Schoehn *et al.*, 2006; Durá *et al.*, 2009). These studies have shown that the overall quaternary structure and internal organization of these three particles are very similar (Franzetti and Durá, 2013). Moreover, they are all strict aminopeptidases with great stability at elevated temperature and  $\text{Co}^{2+}$  ions have a clear stimulatory effect on their amidolytic activities (Durá *et al.*, 2005; 2009; Franzetti and Durá, 2013). However, each PhTET possesses marked substrate specificity: PhTET1 is a glutamyl or aspartyl aminopeptidase with narrow specificity, PhTET2 is a leucyl aminopeptidase that hydrolyses preferentially neutral and aliphatic residues, while PhTET3 is a lysyl aminopeptidase specialized in the removal of large and charged residues from the N-termini of polypeptides (Durá *et al.*, 2005; 2009; Franzetti and Durá, 2013). Because of their complementary action, it has been proposed that the three PhTET peptidase complexes represent an integrated system dedicated to the destruction of a vast variety of intracellular polypeptides (Durá *et al.*, 2009). In this article, we describe the study of their respective oligomeric states *in vivo*. Our results demonstrate that PhTET2 and PhTET3 can interact physically to

form stable multi-subunit peptidasomes capable of breaking down very diverse polypeptidic chains.

## Results

### *The PhTET1, PhTET2, and PhTET3 proteins exist in different oligomeric states in vivo*

We have recently shown that a significant population of the PhTET2 dimeric precursor coexists *in vivo* with the assembled PhTET2 dodecamer (Appolaire *et al.*, 2013). In addition, we have shown that the oligomerization represents a way to trigger the activity of the peptidase towards long peptides (Appolaire *et al.*, 2013). These observations prompted us to examine the oligomeric states of the three PhTETs in *P. horikoshii* cells. For this, post-membrane protein cytoplasmic extracts (S30) were analysed by using continuous sucrose density gradients. The different PhTET proteins were immunodetected in the gradient fractions using specific antibodies. Purified recombinant PhTET dimers and dodecamers were used to calibrate the experiment. The result, presented in Fig. 1, shows that the accumulation of dimeric precursors is a common feature for the three types of PhTET peptidasomes. Moreover, the sedimentation profiles of PhTET2 and PhTET3 were similar, revealing the coexistence of dimeric populations together with assembled dodecamers. Interestingly, in the case of PhTET1, only the dimeric population was detected in the analysed samples. These results suggest a common regulatory mechanism for PhTET2 and PhTET3, and that the assembling of TET1 within the cells obeys a different oligomerization control. It also suggests that PhTET2 and PhTET3 might interact physically to form multi-subunit complexes.



**Fig. 2.** Purification of wild-type recombinant PhTET2:PhTET3 hetero-oligomers coexpressed in *E. coli*.

A. Fractions from the first IEC column (Resource Q) during the purification of recombinant homo-oligomeric PhTET2, PhTET3 and coexpressed PhTET2:PhTET3 loaded on 12% SDS-PAGE gels. The fractions numbers are indicated on the top. The elution position of the hetero-oligomer is between that of homo-oligomeric PhTET2 and PhTET3.

B. Separation of hetero-oligomeric complexes with different subunit composition by successive IEC. Thin lines: elution profile of homomeric PhTET2 (left side) and PhTET3 (right side) proteins. Thick lines: elution profile of recombinant hetero-oligomeric complexes of PhTET2:PhTET3. Top chromatogram: first IEC column (Resource Q) elution profile of the coexpressed hetero-oligomeric PhTET2: PhTET3. The elution volume of the hetero-oligomer is comprised between those of homo-oligomeric PhTET2 and PhTET3. Middle chromatogram: second IEC column (Mono Q) elution profile of the fractions 'A' (plain line), 'B' (half dotted line) and 'C' (dotted line) from the first IEC. Each fraction constitutes of multiple peaks. Each peak corresponds to a complex with different external surface charge, and therefore different subunit composition. Bottom chromatogram: third IEC column (Mono Q) elution profile of the fractions 'a' (plain line), 'b' (half dotted line) and 'c' (dotted line) from the second step respectively labelled 'I', 'II' and 'III'. After three successive IEC columns, three different hetero-oligomeric complexes were purified to homogeneity.

C. Immunodetection of recombinant PhTET2 and PhTET3 proteins from purified coexpressed hetero-oligomers. Eight different purified hetero-oligomeric forms were loaded on 12% SDS-PAGE gels and immunodetected using specific antibodies.

PhTET3 dodecamers (Fig. 2A). Since the two proteins display identical migration on SDS-PAGE, we used the anti-PhTET2 and anti-PhTET3 specific antibodies to confirm that the new eluted protein contained both PhTET2 and PhTET3. We also verified that the purified coexpressed TET sample displayed both PhTET2 and PhTET3 specific activities. These observations are consistent with the efficient formation of at least one hetero-oligomeric TET complex containing PhTET2 and PhTET3 subunits.

The fact that TET complexes are dodecameric particles raises the question of the composition of the purified hetero-oligomeric samples. A series of Mono Q chromatographies was used to further discriminate the different TET species as a function of their subunit compositions. Figure 2B shows the purification of three different TET populations after three steps of IEC runs. After such analyses, 11 different hetero-oligomeric forms were detected. The building block of the dodecameric particle is the dimer (Appolaire *et al.*, 2013). Therefore, if only homo-dimers are considered, no more than seven forms would be possible. Here, eight different forms of TET hetero-oligomers were purified to homogeneity in enough amounts to be characterized. Using anti-PhTET2 and anti-PhTET3 specific antibodies, we confirmed that each of the 8 samples contained both PhTET2 and PhTET3 proteins (Fig. 2C) and the corresponding activities were detected (not shown). This result indicates that, in addition to homo-dimers, hetero-dimers can be formed when PhTET2 and PhTET3 are coexpressed in *E. coli*.

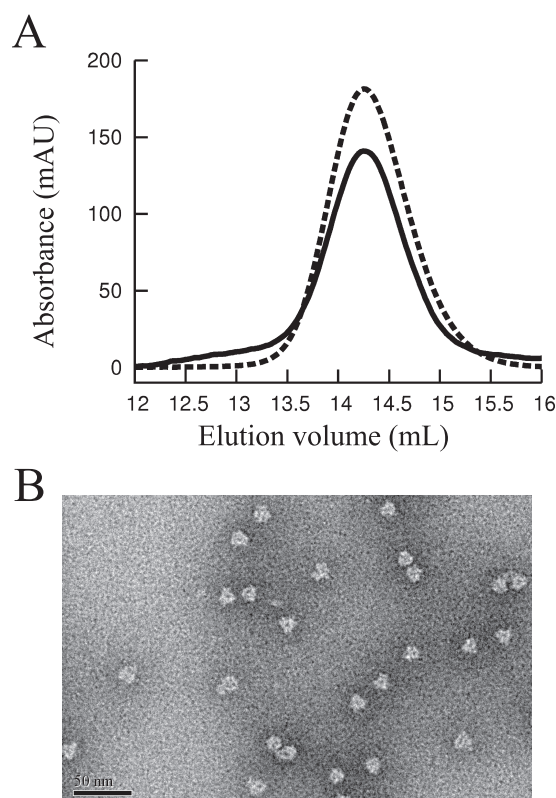
#### The PhTET2 and PhTET3 proteins can form hetero-oligomeric dodecameric complexes after coexpression in *Escherichia coli*

In order to test if the PhTET2 and PhTET3 proteins can form hetero-oligomeric complexes *in vivo*, the wild-type genes were coexpressed in *Escherichia coli* from a pETDuet vector. The experiment yielded an abundant production of soluble TET proteins after cell lysis and heat shock purification. Gel filtration analyses revealed that the coexpressed proteins were assembled as dodecameric particles. The PhTET2 and PhTET3 homo-oligomers display different external surface charge properties. This allowed us to design ion exchange chromatographic (IEC) conditions leading to the clear separation of the two types of complexes. In the same conditions, the major part of the coexpressed TET proteins eluted in between the peaks corresponding to the homo-oligomeric PhTET2 and

Due to their close molecular weights, the PhTET2:PhTET3 stoichiometry of the purified samples could not be accurately determined by using mass spectrometry or quantitative immunoblots. Nevertheless, the stoichiometry of one sample could be assessed based on the properties of *in vitro* reconstituted hetero-oligomers that were produced in order to study the TET oligomerization process. In these studies, the hetero-oligomeric TET particles containing PhTET2 and PhTET3 were obtained by mixing the dimeric precursors recovered after EDTA and basic pH treatment in a re-assembling buffer containing cobalt (Appolaire *et al.*, 2014). The different hetero-oligomeric forms were purified to homogeneity in enough amounts to be analysed by small angle neutron scattering (SANS). This technique allowed the characterization of a particle composed of six subunits of PhTET2 and six subunits of PhTET3 (Appolaire *et al.*, 2014). Analysed on the Mono Q column, this purified complex eluted at the very same conductivity ( $35.9 \text{ mS cm}^{-1}$ ) that one of the assemblies purified to homogeneity after coexpression in *E. coli* (see Fig. 2B, peak II). Therefore, it can be assumed that this complex corresponds to an iso-stoichiometric form of the hetero-oligomer PhTET2:PhTET3. This particle was used to further characterize the biophysical and biochemical properties of the hetero-oligomer. For that, the purified iso-stoichiometric PhTET2:PhTET3 complex was loaded on a size exclusion chromatography column (Superose 6). Its elution volume was identical to that of purified dodecameric PhTET2 protein (Fig. 3A). Furthermore, electron microscopy images of this sample showed homogeneous populations of hollow tetrahedral particles identical to the typical dodecameric TET complexes (Fig. 3B). Thus, the coexpression of PhTET2 and PhTET3 in *E. coli* led to the formation of hetero-oligomeric complexes *in vivo* with the characteristic TET conformation.

*The PhTET2 and PhTET3 proteins assemble as hetero-oligomeric TET dodecameric complexes in P. horikoshii*

In order to assess for the biological relevance of the PhTET2:PhTET3 hetero-oligomers, the properties of native dodecameric TET samples from *P. horikoshii* were further studied. The sucrose density gradient fractions in which TET dodecamers were detected were pooled and analysed in an IEC column (Resource Q). Anti-PhTET1, -PhTET2 and -PhTET3 were used to detect the proteins in the chromatographic fractions (Fig. 4). The experiment showed that most of the native PhTET2 and PhTET3 proteins that are present in the *P. horikoshii* extracts elute with conductivity values that correspond to those of the hetero-oligomeric PhTET2:PhTET3 complexes obtained after coexpression in *E. coli*. However, while most of the PhTET2 and PhTET3 proteins are forming hetero-



**Fig. 3.** Characterization of hetero-oligomers as TET dodecameric complexes.

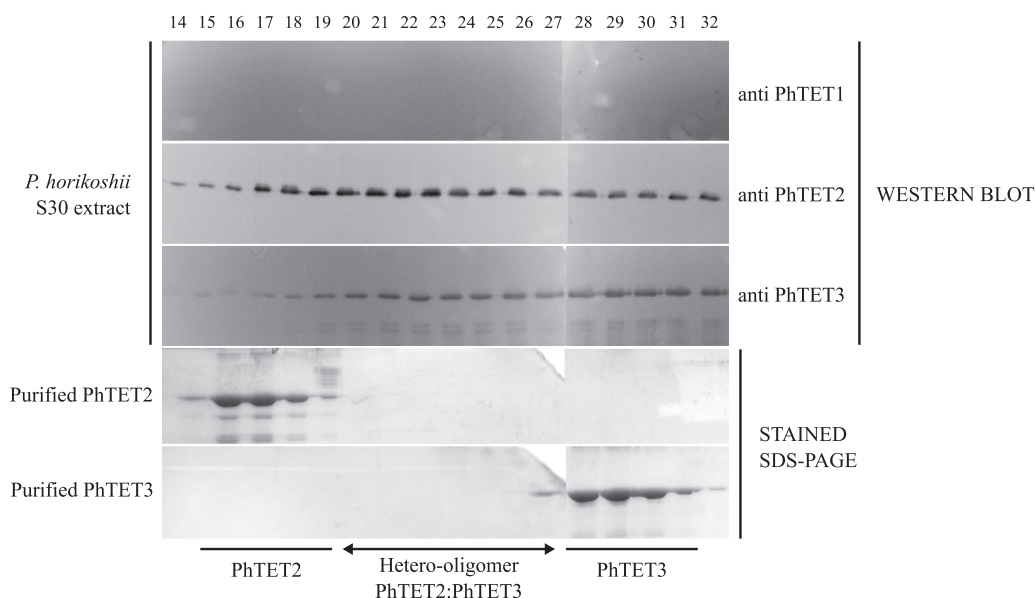
A. Elution profile of a purified recombinant hetero-oligomer (full line) on a size exclusion column (Superose 6) compared to that of the wild-type purified dodecameric PhTET2 protein (dotted line). B. Electron micrograph of the recombinant hetero-oligomer. The picture shows a homogeneous population of hollow tetrahedral-shaped particles.

oligomers, there is still a minor population of PhTET2 and PhTET3 homo-oligomers. In contrast, we noted the absence of PhTET1 in the dodecameric fractions. These data confirm that, *in vivo*, PhTET2 and PhTET3 are naturally designed to assemble as hetero-oligomeric complexes with different stoichiometries.

*The PhTET2:PhTET3 iso-stoichiometric complex displays enhanced polypeptide degradation properties*

For determining the functional benefit of hetero-oligomerization, the catalytic activity of the iso-stoichiometric form of the recombinant hetero-oligomeric PhTET2:PhTET3 complex was compared to that of an equimolar mixture of recombinant homo-oligomeric PhTET2 and PhTET3. Having into account the different substrate specificities of PhTET2 and PhTET3 (Durá *et al.*, 2005; 2009), we used the chromogenic substrate Val-Leu-Lys-pNA in order to compare the catalytic activity of the hetero-complex to that of the mix of homo-complexes. The



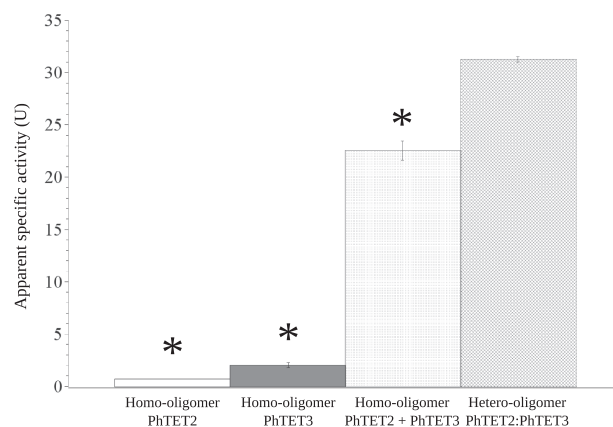


**Fig. 4.** Analytical ion exchange chromatography study of native and recombinant PhTET complexes.

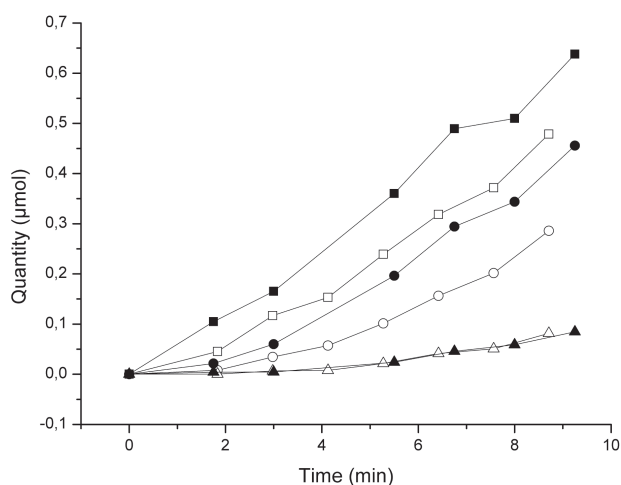
Top panel. Immunodetection of PhTET1, PhTET2 and PhTET3 proteins from *P. horikoshii* cell extracts. The fractions numbers are indicated on the top. The sucrose density gradient high-molecular-weight fractions containing TET dodecamers (7–14, Fig. 1) were pooled and the proteins were separated on a resource Q ion exchange chromatography column. The proteins from fractions 14 to 32 were loaded on 12% SDS-PAGE gels and the different PhTET proteins were immunodetected using specific antibodies.

Bottom panel. Coomassie blue-stained 12% SDS-PAGE. Purified recombinant PhTET2 and PhTET3 homomeric complexes were separated on the resource Q column using the same experimental conditions than the one used for native complexes. A significant part of the native TET complexes eluted in between the homo-oligomeric complexes in the fractions where the iso-stoichiometric PhTET2:PhTET3 hetero-oligomer was detected (indicated by the double arrow). This experiment showed that, in *P. horikoshii*, most of the PhTET2 and PhTET3 proteins are assembled as hetero-oligomers.

formation of coloured para-nitroaniline (pNA) only occurs if the three residues are hydrolysed, and that requires both PhTET2 and PhTET3 activities. Figure 5 shows the almost absence of pNA cleavage in the presence of PhTET2 or PhTET3 alone. In contrast, the combination of PhTET2 and PhTET3 activities in the form of homo-dodecamers allowed the hydrolysis of the peptide, thus showing that the two peptidases can work in concert to degrade polypeptides. However, when the experiment was performed with the purified iso-stoichiometric hetero PhTET2:PhTET3 complex, the apparent specific activity was found to be significantly higher. In order to better understand this behaviour, the time-course of amino acids release resulting from the hydrolytic activities of the different type of TET particles was followed (Fig. 6). This experiment showed that the generation rate of free valine or leucine residues in the reaction mixture is higher when the peptide is processed by the hetero-oligomer. The numbers of PhTET2 and PhTET3 active sites are equivalent in the two different mixtures (equimolar mix of homo-oligomers or iso-stoichiometric form of hetero-oligomer), thus, the co-assembly of PhTET2 and PhTET3 monomers within a single TET particle significantly enhanced the catalytic activity of the system. This demonstrates that the hetero-oligomer possesses both PhTET2 and PhTET3



**Fig. 5.** Aminopeptidase assay using Val-Leu-Lys-pNA as a substrate. White bar: recombinant homomeric PhTET2, grey bar: recombinant homomeric PhTET3, dotted bar: equimolar mixture of homo-oligomeric recombinant PhTET2 and PhTET3, crossed line bar: iso-stoichiometric recombinant PhTET2:PhTET3 hetero-oligomeric complex. The PhTET2:PhTET3 hetero-oligomer shows a significant increased apparent activity compared to an equimolar mixture of homo-oligomeric PhTET2 and PhTET3 against the peptidic substrate ( $*P < 0.01$  Student's *t*-test). The units of apparent specific activity (U) are defined as  $\mu\text{mol}$  of released pNA  $\text{min}^{-1} \text{mg}^{-1}$  of TET.



**Fig. 6.** Time-course of hydrolysis of Val-Leu-Lys-pNA by the mix of homo-oligomers and the hetero-oligomers. Generation of free amino acids in the reaction mixture during the hydrolysis of Val-Leu-Lys-pNA substrate by a mix of recombinant homo-oligomers (□, Val; ○, Leu; △, Lys) and the iso-stoichiometric form of the recombinant hetero-oligomer (■, Val; ●, Leu; ▲, Lys).

activities and an enhanced efficiency to completely degrade peptides compared to an equivalent mixture of homo-oligomers.

In order to assess for the biological significance of this finding, the efficiency of the PhTET hetero-oligomer on a complex mixture of longer peptides was studied. In these experiments, a casein hydrolysate (N-Z-Case – Sigma) was incubated with the different recombinant PhTET particles. The digested peptides were resolved by reverse-phase HPLC and their elution profiles were compared (Fig. 7). This experiment revealed that only a few peptides were digested by PhTET3 while PhTET2 displayed a broader hydrolysis spectrum. When the two homo-dodecameric enzymes were used together, a higher number of peptides were processed. However, the use of the purified iso-stoichiometric hetero-oligomeric PhTET2:PhTET3 complex strikingly enhanced the ability of the enzyme to hydrolyse the mixture of polypeptides. A comparison of the chromatograms shows that some peptides that were only partially hydrolysed by the mix of homo-oligomers were almost fully processed by the hetero-oligomer. Furthermore, the hetero-oligomer displayed activity on peptides that were only poorly hydrolysed, or not processed at all, by the mix of homo-oligomers (Fig. 7). During the experiment, new peaks appeared

due to the production of peptidic products as a result of the amidolytic reactions. The higher amount of these products generated by the hetero-oligomer compared to that of homo-oligomers, alone or in mixture, confirms its enhanced amidolytic activity. In conclusion, these results clearly demonstrate that the hetero-oligomer PhTET2:PhTET3 represent a peptidasome particle with multiple catalytic activities and enhanced efficiency compared to the homo-complexes.

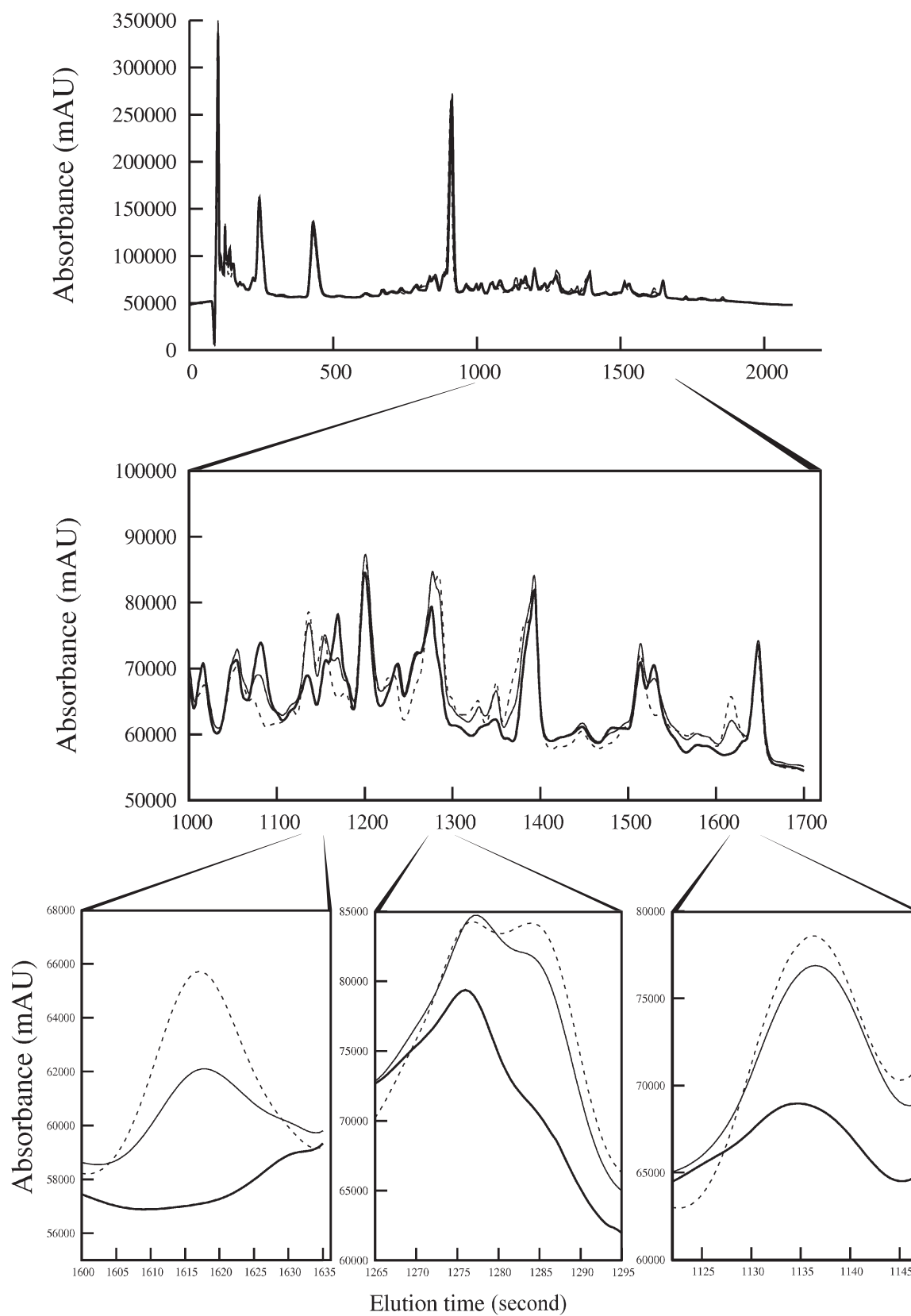
#### Structural and sequence analysis of PhTET

In order to gather insights on why PhTET2 and PhTET3 are specifically able to form hetero-oligomers whereas TET1 is not, a structural and sequence comparison was performed (Fig. 8A). Structural studies of the of *P. horikoshii* TETs have revealed multiple interactions between the dimers that allow the peptidases to assemble as dodecamers. It has been also shown that point mutations in these zones destabilize the dodecameric assemblies. In PhTET2, the R217, R220 and F224 residues are important for the dodecamerization of the particle (Appolaire *et al.*, 2013). The area in which they are contains most of the contacts between dimers. It is situated near the exit hole at the apices of the tetrahedron and the structure of the  $\alpha 5$  helix (following the PhTET2 nomenclature of the PDB code 1Y0R) containing these residues is well conserved between the different PhTETs (Fig. 8A). However, the degree of sequence conservation of this helix, and particularly regarding these three residues, is not equivalent among the three PhTETs. In fact, residues 217, 220 and 224 are very well conserved between TET2 and TET3, but less in TET1 (Fig. 8A).

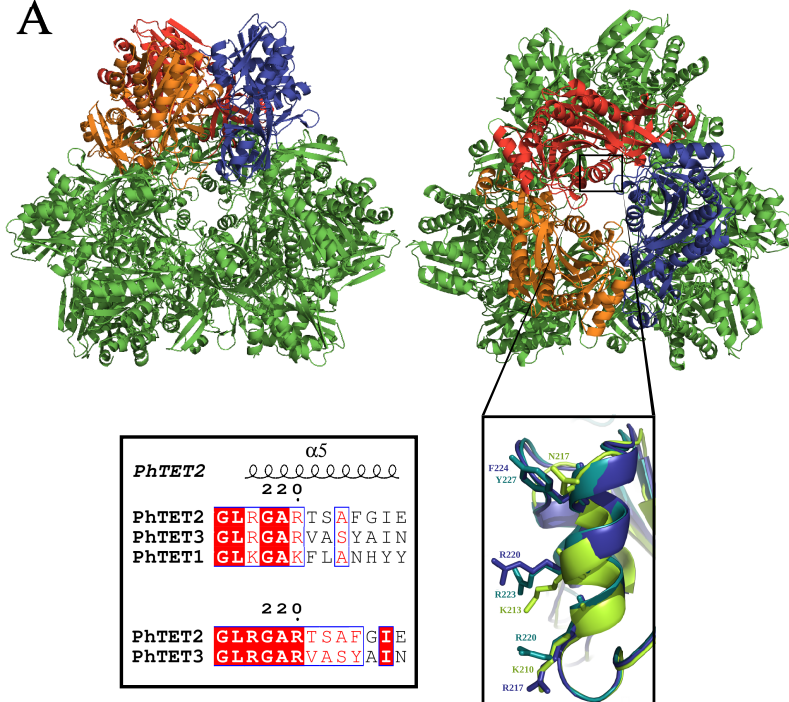
PhTET2 and PhTET3 have overall 48% similarity, while PhTET1 have only 36% and 31% homology with PhTET2 and PhTET3 respectively. Phylogenetic analysis confirms that, although all three TET were inherited in *Pyrococcus* from the ancestor of Thermococcales, TET2 and TET3 are more closely related to each other than they are to TET1 (Fig. 8B). Indeed, a larger phylogenetic analysis including additional TET homologues (not shown) indicates that TET2 and TET3 arose from specific gene duplication in the ancestor of Thermococcales, whereas TET1 has an independent origin. TET2 and TET3 may have therefore diverged progressively to develop the capacity of assembling into hetero-oligomers and, in this way, increased their efficiency, probably related with the general processing of

**Fig. 7.** Comparisons of the peptidase activities of homo-oligomers and the hetero-oligomers on a complex peptide mixture. Chromatograms from reverse phase HPLC experiments. A casein hydrolysate containing peptides of different sizes and compositions was incubated with either an equimolar mixture of homo-oligomeric recombinant PhTET2 and PhTET3 complexes (thin lines) or the iso-stoichiometric recombinant hetero-oligomer PhTET2:PhTET3 (thick lines). The dotted lines correspond to the control experiment. The digestions products were then analysed on a  $\mu$ RPC C2/C18 ST column. Top panel: complete chromatogram of the experiment. Middle and bottom panels: magnification of the chromatogram area highlighting peptides that are hydrolysed more efficiently by the hetero-oligomer than by the mix of homo-oligomers.





A



**Fig. 8.** Comparison of the dodecamerization interface and phylogenetic tree of PhTET1, PhTET2 and PhTET3.

A. View from the face (left hand side) and from the apex (right hand side) of the crystallographic structure of PhTET2 (PDB code: 1YOR). The three monomers forming the catalytic pocket are highlighted in different colour. Below the sequence alignment of the  $\alpha 5$  helix (PhTET2 nomenclature, PDB code: 1YOR) of PhTET1, PhTET2 and PhTET3 is shown. The numbering is that of the PhTET2 protein. This helix is better conserved between PhTET2 and PhTET3 than between PhTET1 and PhTET2 or PhTET3. Close-up: superimposition of the  $\alpha 5$  helix of PhTET1 (green), PhTET2 (light blue) and PhTET3 (dark blue) involved in the dodecamerization interface. The residues mainly implicated in the interactions are represented by sticks.

B. Unrooted phylogenetic tree of TET1, TET2 and TET3 from Thermococcales.

B



peptides, while TET1 may have a different physiological role.

## Discussion

It is generally assumed that in prokaryotes multi-protein complexes tend to have a simpler composition than in eukaryotes (Bochtler *et al.*, 1999; Dacks and Field, 2007; Klinge *et al.*, 2012). We show here that this is not case for the TET system. After coexpression in *E. coli*, and also within *P. horikoshii* cells, two peptidases from the M42 family form hetero dodecameric particles with the typical tetrahedral conformation and activity of the TET aminopeptidases. Many large peptidase assemblies acting on different cellular functions coexist in the cytosol (Groll *et al.*, 2005). Among them, hetero-oligomerization has been observed only in few cases. The eukaryotic 20S proteasome core particle combines three activity types from different beta subunits, thus allowing a more efficient on protein hydrolysis. In the case of the eukaryotic meprins  $\alpha$  and  $\beta$ , hetero-oligomerization occurs to control the membrane binding state of the enzymatic complex (Bertenshaw *et al.*, 2003). The Deg/HtrA peptidase assembly, which has been proposed to play a role in the repair of the damaged photosystem II, can form hetero hexamers made of Deg 5 and Deg 8 subunits in the thylakoid lumen (Sun *et al.*, 2007). The structure of a yeast membrane-bound hetero-oligomeric protein made of the FtsH proteases Yta10 and Yta12, which is involved in the maintenance of mitochondrial functions (Arlt *et al.*, 1998), has been resolved recently (Lee *et al.*, 2011). In the case of TET, our enzymatic assays reveal that hetero-oligomerization represents a way to combine the PhTET2 and PhTET3 catalytic activities in a single cellular nanocompartment, the consequence being an enhanced efficiency compared to the homo-oligomeric complexes. Only a few homo-oligomers were detected *in vivo* after coexpression both in *E. coli* (Fig. 2) and in *P. horikoshii* (Fig. 4) cellular extracts. This indicates that the formation of hetero-oligomers is a very efficient process and that the PhTET2:PhTET3 particles represent peptidasomes *in vivo*. To our knowledge, this is the first report for such a mechanism aimed to improve the multi-enzymatic assembly efficiency among exopeptidases.

Our results demonstrate that the hetero-oligomers are more efficient at hydrolysing complex polypeptidic mixtures when compared to the homo-oligomers. TET peptidases are non-processive enzymes, that is to say, they release the peptide substrate from the catalytic site after cleavage of the N-terminal amino acid (Durá *et al.*, 2005; 2009; Franzetti and Durá, 2013). Therefore, the presence of two complementary active sites within a single cellular nanocompartment represent a clear advantage compared to the situation in which the PhTET2 and PhTET3 homo-

oligomers would act in concert, and thus, where the polypeptidic product would have to exit one of the particles to be further processed by the other enzymatic complex. In addition, we have recently demonstrated that the TET peptidase oligomerization process is not random and involves the formation of hexameric intermediates (Appolaire *et al.*, 2013). In the case of the *in vitro* reconstructed PhTET2:PhTET3 dodecameric particles, the use of SANS and contrast variation revealed that this controlled oligomerization process results in the obligatory formation of catalytic chambers made of the combination of the two different active sites (Appolaire *et al.*, 2014). Thus, after cleavage, the free N-terminus of the polypeptidic product can be easily redirected towards the most suitable active site with respect to the identity of its P1 amino acid. The clear and abundant formation of hetero-oligomers *in vivo* that we report here also suggests that the evolution has modified the inter-subunit contacts in order to favour the formation of PhTET2:PhTET3 peptidasomes within cells.

It is not uncommon that new functions arise from gene duplication followed by evolutionary adaptation (Ohno, 1970). Moreover, the products of the duplicated genes often maintain a functional linkage, either as enzymes acting in tandem (Rigas *et al.*, 2012) or as integrated components of multi-subunit complexes (Hughes, 1997). Following specific gene duplication in the ancestor of Thermococcales, PhTET2 and PhTET3 likely diverged to give rise to two types of TET proteins with marked substrate specificities and improvement of the TET system performances as observed in our study, likely linked to adaptation to the peptide-based energy metabolism of the deep-sea thermococcales (Adams *et al.*, 2001). In contrast, the independent origin of PhTET1 is consistent with the fact that it does not interact physically with PhTET2 and PhTET3. In fact PhTET1 does not possess the same inter-subunit contacts as those used in the PhTET2 and PhTET3 edifices. Therefore, our study indicates that evolution following gene duplication may also act on the oligomerization state of specific cellular machineries.

In a previous study, we have demonstrated that oligomerization triggers the enzymatic activity of PhTET2 towards large peptides and that the dimeric precursor of this enzyme accumulates in *P. horikoshii* cells (Appolaire *et al.*, 2013). Here, we reveal that this last feature is also true for the two other PhTET peptidases. This suggests that, as a first step, there is a common mechanism of regulation for all PhTETs in which a dimeric precursor accumulates in the cell before being activated through oligomerization. The signal that stimulates this remains, however, to be characterized. But interestingly, our results show, moreover, that the cellular PhTET1 oligomerization state appears to be very different from those of PhTET2 and PhTET3. PhTET1 possesses a very narrow substrate specificity (Franzetti and Durá, 2013). This, together with

the *in vivo* observations, suggests that PhTET1 displays a different physiological role than the PhTET2:PhTET3 peptidasomes, which appear to be aimed at destroying polypeptidic chains in an unspecific manner. Two oligomeric forms of PhTET1 have been described *in vitro* (Schoehn *et al.*, 2006): the protein is able to generate tetrahedral dodecameric particles, and also large octahedral tetracosameric complexes. Therefore, the oligomerization of its dimers might be triggered by factors absent in our experimental conditions, such as some type of stress or the ageing of the cells. Conversely, the function of the hetero-PhTET2:PhTET3 may be essential for metabolism and/or to work in concert with the proteasome. Finally, due to its enhanced proteolytic efficiency against complex substrates, the multi-subunit TET peptidasome could be employed in biotechnological and industrial processes such as the production of bioactive compounds from dairy products.

## Experimental procedures

### *Cloning and coexpression of P. horikoshii TET genes in E. coli*

The PCR fragments of wild-type PhTET1, PhTET2 and PhTET3 genes (*PH0519*, *PH1527*, and *PH1821* respectively) were double digested and cloned into the BamHI–HindIII restriction sites of pETDuet to generate the recombinant plasmids pETDuet-TET1-TET2; pETDuet-TET1-TET3 and pETDuet-TET2-TET3. The recombinant plasmids were transformed into *E. coli*. The cells were grown in 1 l of LB medium containing ampicillin (50 µg ml<sup>-1</sup>). After incubation with shaking at 37°C until the A<sub>600</sub> reached 0.6–1.0, the induction was carried out by adding isopropyl β-D-thiogalactopyranoside at a final concentration of 0.05 mM and shaking for 4 h at 37°C. The induced cells were harvested by centrifugation. The cell pellets were suspended in 50 ml of 50 mM Tris-HCl, 150 mM NaCl, 0.1% Triton X-100, pH 8.0, and maintained at –80°C until used.

### *P. horikoshii cells cultivation and sucrose density gradient fractionation*

*Pyrococcus horikoshii* cells growth, post-membrane protein cytoplasmic extracts (S30) preparation, and protein fractionation by sucrose density gradients were performed as described in Appolaire *et al.* (2013). The proteins separated by sucrose density gradient fractionations were precipitated prior to SDS-PAGE analysis as described in Appolaire *et al.* (2013).

### *Purification of recombinant homo-oligomeric and hetero-oligomeric TET complexes*

Cell suspensions were thawed at room temperature and the following added: 12.5 mg of lysozyme (Euromedex), 2.5 mg of DNase I grade II (Roche), 10 mg of RNase (Roche), 50 mg

of Pefabloc SC (Roche) and 0.5 ml of 2 M MgSO<sub>4</sub>. The disruption of the cells was achieved by sonication in a Branson sonifier 150 at 4°C. Five 30 s bursts at intensity 10 with intermediary pauses of 30 s were employed. The crude extract was then heated at 85°C for 15 min to eliminate most mesophilic proteins from the *E. coli* host, the lysates were clarified by centrifugation at 17 400 g for 1 h and, after adjusting its Tris concentration to 20 mM and its pH to 7.5, the supernatant was loaded on a 6 ml Resource Q column (GE Healthcare) equilibrated in 150 mM NaCl, 20 mM Tris-HCl, pH 7.5. After washing with three column volumes (CV) of this buffer, bound proteins were eluted at 4 ml min<sup>-1</sup> with a 20 CV linear salt gradient (0.15–0.45 M NaCl in 20 mM Tris-HCl, pH 7.5). The same procedure was used for homo-oligomeric and hetero-oligomeric TET complexes.

In order to separate the different hetero-oligomeric complexes as a function of conductivity, and therefore as a function of their subunit composition, the hetero-oligomeric fractions from the Resource Q column were combined in eight pools, the salt concentration of the samples was adjusted to 250 mM NaCl, and then, they were further resolved on a 1 ml Mono Q column (GE Healthcare) equilibrated in 250 mM NaCl, 20 mM Tris-HCl, pH 7.5. After washing with three column volumes of this buffer, bound proteins were eluted at 1 ml min<sup>-1</sup> with a 20 CV linear salt gradient (0.25–0.45 M NaCl in 20 mM Tris-HCl, pH 7.5). Multiple peak chromatograms were obtained. The fractions corresponding to identified peaks were pooled, the buffer was adjusted to 0.25 M NaCl, and the pools loaded again on the same Mono Q column and in the same conditions, until elution of symmetrical peaks corresponding to the purified different hetero-oligomeric PhTET species was achieved. The purified complexes were analysed thereafter by size exclusion on a Superose 6 column (GE Healthcare) equilibrated and run in 150 mM NaCl, 20 mM Tris-HCl, pH 7.5 to control their final oligomerization state.

### *Protein immunodetection*

The PhTET1, PhTET2 and PhTET3 proteins were immunodetected in sucrose density gradient and ion exchange chromatography fractions. One volume of SDS-PAGE loading buffer (50 mM Tris-HCl, 8 M urea, 2 M thiourea, 100 mM DTT, 3% SDS, 0.1% bromophenol blue, 10% glycerol, pH 6.8) was then added to the samples. These were heated at 95°C for 7 min and their proteins were separated by SDS-PAGE using 12% gels.

The proteins were transferred on Hybond-P PVDF-membranes, and immunoreactive bands were visualized by chemiluminescence as detailed by the supplier (ECL detection kit; GE Healthcare). The signal was detected using a Kodak Image station 4000 mm. Specific antibodies were raised against the three different PhTET proteins. For this, peptides were chosen in the protein primary sequences, synthesized and injected to rabbits. The anti-PhTET2 antibody was raised against DERDVDATVELMTKALENIHELKI as described (Appolaire *et al.*, 2013). The anti-PhTET1 and -PhTET3 antibodies were raised against MMSMIEKLKFTQ and EAARQLGDHEADYIVGS by Eurogentec and Proteogenix respectively. The antisera were used at 1:6000, 1:2500 and 1:5000 dilutions for PhTET1, PhTET2 and PhTET3 respectively.



## Electron microscopy

In order to control the conformation of the hetero-oligomeric PhTET complexes, 4  $\mu\text{l}$  of purified sample ( $\sim 0.1 \text{ mg ml}^{-1}$ ) were loaded between mica-carbon interfaces as described in Franzetti *et al.* (2002). The sample was stained using 2% sodium silicotungstate, pH 7.5, and air-dried. Images were taken under low-dose conditions in a CM12 Philips electron microscope working at 120 kV and with a nominal magnification of 40 000 using an Orius SC1000 CCD camera.

## Enzyme activity assays

**Assay on amino-acyl-para-nitroanilide.** Hydrolytic activity on different synthetic chromogenic compounds was determined using aminoacyl-pNAs through the following standard procedure. Reactions were initiated by addition of 3  $\mu\text{l}$  of enzyme (0.3 or 1.5  $\mu\text{g ml}^{-1}$ ) to a 57  $\mu\text{l}$  pre-warmed mixture containing the chromogenic compound (5 mM) in 50 mM HEPES, 150 mM KCl, 1 mM  $\text{CoCl}_2$ , pH 7.5, and covered with mineral oil to avoid water evaporation. Incubations were performed at 85°C for 5 or 10 min, and the reactions were stopped by adding 60  $\mu\text{l}$  of 0.2 M acetic acid and cooling at 0°C. The quantity of pNA liberated was measured at 405 nm in a VICTOR 1420 multilabel counter (Wallac) or in a Synergy H4 hybrid reader (Biotek). Four replicates and four enzyme blanks were assayed for each experimental point. Statistical analysis was performed using Origin Pro v 8.5.

For the study of Val-Leu-Lys-pNA degradation, the enzyme (3  $\mu\text{g ml}^{-1}$ ) was incubated with the substrate (5 mM) using the standard assay conditions. Incubations were performed at 75°C, at determined times, 80  $\mu\text{l}$  aliquots were removed from the reaction mixture and added to 200  $\mu\text{l}$  of cold acetonitrile in order to stop the reaction and precipitate the protein (Aristoy and Toldrà, 1991). Samples were maintained at 4°C for 15 min, and the protein was removed by centrifugation. Fifty microlitres of supernatant were dried under vacuum and, after resuspension of the solids in analysis buffer, quantitative amino acid analysis was performed. Amino acids were analysed on a model Biochrom 30 amino acid analyser, using the standard sodium citrate eluting buffer system. Calibrations were made with standard solutions which contained all the amino acids except tryptophan.

**Assay on N-Z-CASEIN hydrolysate.** The different TET peptidases (2.5  $\mu\text{g ml}^{-1}$ ) were incubated with a casein enzymatic hydrolysate (N-Z-Case – Sigma) at 10  $\text{mg ml}^{-1}$  in the standard assay conditions. Incubations were performed at 85°C during 30 min. The reactions were stopped by cooling the sample at 0°C. Eighty microlitres of each sample were added to 720  $\mu\text{l}$  of diluent for the chromatographic separation (phase A, see later) and kept at 4°C until analysis. Samples were centrifuged for 10 min at 10 000  $g$  prior to the injection. Fifty microlitres were injected on a  $\mu\text{RPC C2/C18 ST}$  column (4.6  $\times$  100 mm) (Amersham) in a Perkin Elmer HPLC system. Phase A was composed of trifluoroacetic acid at 0.065% in acetonitrile at 2% in water, and phase B contained trifluoroacetic acid at 0.05% in acetonitrile at 80% in water. Bound peptides were eluted at 0.8  $\text{ml min}^{-1}$  with a 20 ml linear gradient (0–40% phase B), and were detected by measuring their absorbance at 214 nm. Peaks were detected and compared using TotalChrom v. 6.3.1 (Perkin Elmer).

## Structural and sequence analysis

In order to compare their structure, the crystallographic model of the PhTET1, PhTET2 and PhTET3 dimers (PDB code 2CF4, 1Y0R and 2WZN) were superimposed using the program COOT. Illustrations were produced using the program Pymol.

Multiple alignment of the three PhTETs was obtained by Muscle (Edgar, 2004) with default parameters. The alignment was visualized by using the suite Seaview (version 4.2.5) (Gouy *et al.*, 2010).

Representation of the alignment excerpt was generated using ESPript (<http://esprict.ibcp.fr>).

The multiple alignment was trimmed by using BMGE (Criscuolo and Gribaldo, 2010) with default parameters providing a matrix of 319 unambiguously aligned amino acid positions for phylogenetic analysis. Maximum likelihood phylogenetic analysis was performed by using PhyML 3.0, the LG model of amino acid substitution and a gamma correction to take into account heterogeneity of evolutionary rates at sites (Guindon *et al.*, 2010) as implemented in Seaview. Statistical support at nodes was calculated by non-parametric bootstrap on 100 resamples on the original data set.

## Acknowledgements

We acknowledge financial supports from the Agence Nationale de la Recherche (Grant 'Archelyse' ANR-12-BSV8-0019-01). A. A. was supported by a PhD scholarship from the French Ministry for Research and Technology. We thank J. Le Bars for help in *P. horikoshii* cultivation. We thank D. Fenel and C. Moriscot from the PSB/IBS platform for the electron microscopy study. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

## References

- Adams, M.W., Holden, J.F., Menon, A.L., Schut, G.J., Grunden, A.M., Hou, C., *et al.* (2001) Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* **183**: 716–724.
- Appolaire, A., Rosenbaum, E., Dura, M.A., Colombo, M., Marty, V., Savoye, M.N., *et al.* (2013) *Pyrococcus horikoshii* TET2 peptidase assembling process and associated functional regulation. *J Biol Chem* **288**: 22542–22554.
- Appolaire, A., Girard, E., Colombo, M., Durá, M.A., Moulin, M., Härtlein, M., *et al.* (2014) Small angle neutron scattering reveals the assembling mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase. *Acta Crystallogr D Biol Crystallogr* (in press).
- Aristoy, M.C., and Toldrà, F. (1991) Deproteinization techniques for HPLC amino acid analysis in fresh pork muscle and dry-cured ham. *J Agric Food Chem* **39**: 1792–1795.
- Arlt, H., Steglich, G., Perryman, R., Guiard, B., Neupert, W., and Langer, T. (1998) The formation of respiratory chain

- complexes in mitochondria is under the proteolytic control of the m-AAA protease. *EMBO J* **17**: 4837–4847.
- Bertenshaw, G.P., Norcum, M.T., and Bond, J.S. (2003) Structure of homo- and hetero-oligomeric meprin metalloproteases. Dimers, tetramers, and high molecular mass multimers. *J Biol Chem* **278**: 2522–2532.
- Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., and Huber, R. (1999) The proteasome. *Annu Rev Biophys Biomol Struct* **28**: 295–317.
- Criscuolo, A., and Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* **10**: 210.
- Dacks, J.B., and Field, M.C. (2007) Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *J Cell Sci* **120**: 2977–2985.
- Durá, M.A., Receveur-brechot, V., Andrieu, J.-P., Ebel, C., Schoehn, G., Roussel, A., and Franzetti, B. (2005) Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry* **44**: 3477–3486.
- Durá, M.A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F.M.D., and Franzetti, B. (2009) The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol Microbiol* **72**: 26–40.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
- Franzetti, B., and Durá, M.A. (2013) Chapter 371 – PhTET1 aminopeptidase. In *Handbook of Proteolytic Enzymes*. Neil, G.S., and Rawlings, D. (eds). Oxford: Academic Press, pp. 1638–1645.
- Franzetti, B., Schoehn, G., Hernandez, J.F., Jaquinod, M., Ruigrok, R.W., and Zaccari, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO J* **21**: 2132–2138.
- Gonzales, T., and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol Rev* **18**: 319–344.
- Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224.
- Groll, M., Bochtler, M., Brandstetter, H., Clausen, T., and Huber, R. (2005) Molecular machines for protein degradation. *ChemBiochem* **6**: 222–256.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**: 307–321.
- Hughes, A.L. (1997) Evolution of the proteasome components. *Immunogenetics* **46**: 82–92.
- Ingmer, H., and Brondsted, L. (2009) Proteases in bacterial pathogenesis. *Res Microbiol* **160**: 704–710.
- Kim, D., San, B.H., Moh, S.H., Park, H., Kim, D.Y., Lee, S., and Kim, K.K. (2010) Structural basis for the substrate specificity of PepA from *Streptococcus pneumoniae*, a dodecameric tetrahedral protease. *Biochem Biophys Res Commun* **391**: 431–436.
- Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., and Ban, N. (2012) Atomic structures of the eukaryotic ribosome. *Trends Biochem Sci* **37**: 189–198.
- Lee, S., Augustin, S., Tatsuta, T., Gerdes, F., Langer, T., and Tsai, F.T.F. (2011) Electron cryomicroscopy structure of a membrane-anchored mitochondrial AAA protease. *J Biol Chem* **286**: 4404–4411.
- Lutfullah, G., Azhar, N., Amin, F., Khan, Z., Azim, M.K., Shouqat, K., et al. (2009) Structural bioinformatics of *Vibrio cholerae* aminopeptidase A (PepA) monomer. *Protein Pept Lett* **16**: 36–45.
- Ohno, S. (1970) *Evolution by Gene Duplication*. Berlin, New York: Springer-Verlag, p. xv.
- Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **40**: D343–D350.
- Rigas, S., Daras, G., Tsitsekian, D., and Hatzopoulos, P. (2012) The multifaceted role of Lon proteolysis in seedling establishment and maintenance of plant organelle function: living from protein destruction. *Physiol Plant* **145**: 215–223.
- Russo, S., and Baumann, U. (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J Biol Chem* **279**: 51275–51281.
- Schoehn, G., Vellieux, F.M., Asuncion Dura, M., Receveur-Brechot, V., Fabry, C.M., Ruigrok, R.W., et al. (2006) An archaeal peptidase assembles into two different quaternary structures: a tetrahedron and a giant octahedron. *J Biol Chem* **281**: 36327–36337.
- Sun, X., Peng, L., Guo, J., Chi, W., Ma, J., Lu, C., and Zhang, L. (2007) Formation of DEG5 and DEG8 complexes and their involvement in the degradation of photodamaged photosystem II reaction center D1 protein in Arabidopsis. *Plant Cell* **19**: 1347–1361.

### 4.3 DISCUSSION ET PERSPECTIVES

---

Ce travail montre que les protéines PhTET2 et PhTET3 sont les composants d'un même complexe dodecamérique dans les cellules de *P. horikoshii*. Ce complexe montre des propriétés de destruction élargie par rapport aux complexes homo-oligomériques. Des travaux complémentaires réalisés par Matteo Colombo au laboratoire indiquent par ailleurs qu'en présence de cobalt, PhTET3 pourrait éliminer efficacement des acides aspartiques et glutamiques des extrémités N-ter des polypeptides (résultats non publiés). Ainsi, le complexe hétéro-oligomérique TET2-TET3 serait un peptidasome susceptible de dégrader complètement les peptides intracellulaires. Ceci le placerait comme un acteur principal dans la destruction de peptides issus du protéasome ou en provenance du milieu extracellulaire. Dans ces conditions, on peut s'interroger sur le rôle de PhTET1 qui provient lui aussi d'un transfert de gène bactérien. Des analyses en HPLC de profils de digestion peptidiques, non présentées dans l'article montrent que son spectre d'action est réduit par rapport à PhTET2 :PhTET3. De plus, nous montrons ici qu'il n'est pas oligomérisé dans la cellule en même temps que PhTET2 :PhTET3. Le travail présenté dans le chapitre 5 de cette thèse montre que l'oligomérisation représente un moyen pour les peptidases TET d'augmenter leur efficacité de dégradation vis à vis de long polypeptides. Ainsi l'oligomérisation *in vivo* pourrait représenter un mécanisme de régulation de l'activité des peptidases TET. Ceci suggère que PhTET1 n'est pas "activé" dans les cellules en même temps que le peptidasome PhTET2:PhTET3 et qu'il posséderait donc un rôle physiologique différent. Nous avons montré que PhTET1 est spécialisé dans la dégradation d'extrémités N-ter riches en acide aspartique et glutamique. Ces acides aminés n'ont pas besoin d'être modifiés par des amino-transférases pour alimenter le cycle de Krebs et sont donc probablement très précieux en conditions de carence nutritionnelle. Une hypothèse de travail est que l'activité de PhTET1 est mobilisée par la cellule uniquement dans des situations de carence nutritionnelle. Dans ces conditions il serait intéressant d'étudier l'état d'oligomérisation de TET1 et de TET2 :TET3 dans des extraits cellulaires provenant de cellules de *Pyrococcus horikoshii* placées dans différentes situations physiologiques: phase de croissance, stress physico-chimique, carences nutritionnelles, métabolisme hétérotrophe ou autotrophe, utilisation de glucides ou de peptides comme sources d'énergie, inhibition du protéasome, etc... Ces expériences pourraient être complétées



par une étude de l'accumulation des protéines TET dans les cellules et une étude de leur régulation transcriptionnelle par RT-PCR. Un mutant KO d'une des deux peptidases TET de *H. volcanii* a déjà été réalisé dans le passé par Cédric Norais dans l'équipe de H. Myllykalio à Orsay (résultats non publiés). Ce mutant est viable. Des études génétiques chez une *Thermococcale* (inactivation de gènes ou surexpressions hétérologues) associées à des analyses phénotypiques (peptidomique et métabolomique) seraient nécessaires pour valider nos hypothèses et préciser le caractère essentiel ou non des différentes peptidases TET pour la cellule. L'outil génétique est en effet disponible chez *P. furiosus*, *T. kodakarensis* ou *T. barophilus* (Thiel *et col.*, 2014).

### 5. ASSEMBLAGE DES AMINOPEPTIDASES TET : PROCESSUS ET SIGNIFICATION BIOLOGIQUE.

#### 5.1 PRESENTATION DU TRAVAIL:

---

Les expériences de fractionnement cellulaire d'extraits de *Pyrococcus horikoshii* indiquent qu'une grande partie des peptidases TET existent sous une forme d'assemblages de bas poids moléculaires *in vivo*. Nous avons montré ce phénomène par des expériences de fractionnement sur gradient de sucrose identiques à celles décrites dans le chapitre 4. Ces expériences ont été réalisées chez *P. horikoshii*, avec du matériel cellulaire préparé par Anne Godfroy à l'IFREMER à Brest et également chez une archée halophile extrême, *Halobacterium salinarum*, cultivée au laboratoire. Ceci est étonnant car lorsqu'on exprime ces peptidases sous forme recombinante dans *E. coli*, on obtient essentiellement des formes dodécamériques. Les études de stabilités thermiques, de résistance aux pH extrêmes et même de stabilité vis à vis des très hautes pressions hydrostatiques montrent que les peptidases TET, une fois assemblées, sont extraordinairement robustes (Rosenbaum *et col.*, 2012). De même, les expériences de diffusion neutronique présentées dans le chapitre 6 de la partie résultat de cette thèse ont révélé l'absence d'échange entre sous unités des complexes TET. Ceci nous a conduit à nous intéresser à l'activité biologique de ces formes de bas poids moléculaires ainsi qu'aux mécanismes d'assemblage des peptidases TET. Le but de cette étude était aussi de comprendre pourquoi les peptidases TET forment des complexes auto-compartimentés alors que l'assemblage et le fonctionnement des sites actifs des métallopeptidases de la famille M42 ne semblent pas nécessiter la formation d'une structure quaternaire. D'autre part, les TET apparaissent impliquées dans la dégradation de peptides non structurés. On peut donc s'interroger sur les raisons de la formation de structures oligomériques complexes chez des aminopeptidases.

Pour aborder cette question il nous fallait identifier la forme de bas-poids moléculaire accumulée dans les cellules. Les travaux du laboratoire avaient déjà montré une importance des métaux catalytiques dans le processus d'assemblage (Rosenbaum *et col.*, 2011). Cette étude montre qu'il est possible de désassembler les complexes PhTET en incubant les complexes pendant plusieurs jours en présence d'EDTA. La forme

dimérique de TET apparaissait comme une espèce métastable capable de se réassembler de manière séquentielle lorsque des métaux (Co ou Zn) sont rajoutés dans le tampon de ré-oligomérisation. Toutefois cette méthode permet difficilement d'isoler les intermédiaires du repliement et la purification homogène de dimères du fait des long temps de dialyse. Des chocs de pH important sont nécessaires et le processus doit être finement contrôlé pour éviter un excès d'agrégats. De plus, les formes intermédiaires obtenues ne sont pas actives et ne correspondent pas à des formes physiologiques du fait de l'absence de métaux. Métaux qui jouent eux-mêmes un rôle important dans la structure de la protéine.

Avant mon arrivée au laboratoire, M. A Durà et E. Rosenbaum avaient réalisé un important travail de mutagenèse dirigée ciblée sur les résidus susceptibles de moduler les interactions entre les sous unités de la peptidase PhTET2. Celle-ci a été choisie car elle permet d'obtenir les plus hauts rendements de purification. Vingt-quatre mutants simples et multiples avaient été ainsi réalisés. Les protéines mutantes dodécamériques obtenues avaient toute été caractérisées par diffusion aux petits angles (SAXS) pour s'assurer de leur structure quaternaire. Leur thermostabilité ainsi que leur efficacité enzymatique respective avaient aussi été mesurées et comparées à la forme sauvage de la protéine. Ce travail a conduit à retenir un penta-mutant pour notre étude. Nous avons montré que, exprimé dans *E. coli*, la protéine portant les 5 mutations s'accumulait sous la forme d'une espèce de bas poids moléculaire identique à celle détectée dans *Pyrococcus horikoshii*. J'ai mis au point des protocoles de chromatographie permettant de purifier cette forme de bas poids moléculaire. En effet, il est apparu que les formes de différents poids moléculaires de la peptidase TET qui sont produites lorsque le penta-mutant est exprimé chez *E. coli* ont des comportements différents sur colonne échangeuse d'ions. Ce caractère a permis de séparer lors d'une première étape de chromatographie échangeuse d'ions sur colonne ResourceQ les formes de bas poids moléculaires des formes de hauts poids moléculaire tel qu'il est décrit dans l'article qui suit. Une seconde étape de gel filtration permettait alors d'obtenir des échantillons monodisperse des différentes formes du mutant de l'aminopeptidase PhTET2. Cependant, ces dernières n'étaient pas stables dans le temps. En effet, après purification de la forme de plus bas poids moléculaire observée, il apparaissait des formes de plus haut poids moléculaires après quelques heures. Ce résultat suggérait que les formes de

bas poids moléculaires du mutant, une fois purifiées, étaient capables de se réassembler. Bien que très intéressante, cette propriété posait néanmoins un problème majeur pour la caractérisation des différentes formes observées.

J'ai donc identifié des conditions physicochimiques me permettant d'abord de stabiliser cette espèce, puis de contrôler sa cinétique d'assemblage en particules TET dodécamériques en tout point identiques aux formes sauvages. J'ai ainsi montré qu'il s'agissait d'un précurseur de l'assemblage de la particule TET. Les expériences d'ultracentrifugation analytique indiquaient qu'il s'agissait d'un dimère de TET. La bonne maîtrise de la biochimie du système m'a permis de purifier ce précurseur en quantité et en homogénéité suffisante pour réaliser des études SAXS à l'ESRF, à l'époque sur la ligne BM16 avec E. Rosenbaum. Nous avons pu collecter des données seulement 7 minutes après l'élution de la protéine de la dernière colonne de filtration sur gel. Dans toutes ces expériences le plus grand soin a été apporté pour s'assurer que la monodispersité et l'état de la protéine n'avaient pas changé pendant le temps expérimental (contrôles avant et après la mesure par chromatographie d'exclusion et gels natifs). J'ai ainsi pu déterminer la structure de l'enveloppe du dimère précurseur de TET. J'ai également pu suffisamment ralentir le processus d'assemblage du dimère pour suivre l'ordre d'apparition des différents complexes intermédiaires par ultracentrifugation analytique (AUC) et gels natifs. Ces données combinées avec une analyse *in silico* des contacts inter sous-unité dans les structures cristallographiques réalisée avec M. Colombo et Eric Girard ont permis de proposer un mécanisme d'assemblage. L'étude par microscopie électronique des échantillons comportant les différents intermédiaires d'assemblage a été réalisée à l'IBS par G. Schoehn. Elle a permis de confirmer notre modèle. Le travail présenté dans le chapitre 6 de cette thèse permettra de valider expérimentalement le modèle proposé.

La deuxième partie du travail présenté dans ce chapitre a consisté à comprendre les différences fonctionnelles entre les formes assemblées et non-assemblées du dimère de la protéine TET. En effet, la structure en solution du dimère montrait une conformation identique à celle qu'il adopte au sein de la particule dodécamérique assemblée.

Les conditions de stabilité particulières du mutant ont nécessité la mise au point de protocole de purification et de test enzymatique afin que les résultats obtenus soient

significatifs et reproductibles. Les difficultés à purifier le dimère de PhTET2 m'ont d'abord mené à réaliser une série de tests enzymatiques sur gel natif. Cette méthode ne s'est pas révélée suffisamment sensible. J'ai donc choisi une stratégie différente. J'ai optimisé les conditions de purification afin que la concentration d'enzyme en sortie de colonne de gel filtration soit légèrement supérieure aux concentrations nécessaires pour les tests enzymatiques. Cette stratégie m'a permis d'obtenir des échantillons prêts pour analyse quasi immédiatement après la purification, ces derniers ne nécessitant qu'une légère dilution. Les manipulations qui auraient pu influencer sur la stabilité du mutant étaient alors limitées. L'état d'oligomérisation des échantillons était alors contrôlé une nouvelle fois après les tests par gel filtration ou gel natif. Ainsi, durant l'ensemble des tests enzymatiques réalisés sur le mutant, l'état d'oligomérisation de la particule était contrôlé. Les tests enzymatiques sur peptides courts sont les mêmes que ceux utilisés en routine et présentés dans le chapitre 4.

Pour comprendre les différences qu'il existe entre le dimère et le dodécamère, j'ai comparé leurs activités sur des peptides de différentes tailles. J'ai pour cela utilisé un test d'activité mis au point par F. Frottin dans le laboratoire de T. Meinnel. Ce test permet d'extraire rapidement des constantes enzymatiques de réaction où un seul résidu du peptide substrat est hydrolysé. La taille et la nature du peptide substrat n'est pas limitante. Le principe du protocole utilisé est présenté dans le chapitre « Méthodes » (voir section 2.1.2.2). Il a ainsi été possible de montrer une différence significative d'activité entre les formes libres et assemblées du dimère.

**5.2    *PYROCOCCUS HORIKOSHII* TET2 PEPTIDASE ASSEMBLING  
PROCESS AND ASSOCIATED FUNCTIONAL REGULATION**

---

# Pyrococcus horikoshii TET2 Peptidase Assembling Process and Associated Functional Regulation<sup>\*[5]</sup>

Received for publication, January 3, 2013, and in revised form, May 13, 2013. Published, JBC Papers in Press, May 21, 2013, DOI 10.1074/jbc.M113.450189

Alexandre Appolaire<sup>†1</sup>, Eva Rosenbaum<sup>‡</sup>, M. Asunción Durá<sup>‡2</sup>, Matteo Colombo<sup>‡</sup>, Vincent Marty<sup>‡</sup>, Marjolaine Noirclerc Savoye<sup>‡</sup>, Anne Godfroy<sup>§</sup>, Guy Schoehn<sup>‡</sup>, Eric Girard<sup>‡</sup>, Frank Gabel<sup>‡</sup>, and Bruno Franzetti<sup>‡3</sup>

From the <sup>†</sup>Institut de Biologie Structurale, CNRS, UMR5075, F-38027/Commissariat à l'Energie Atomique, F-38054/Université Joseph Fourier, F-38027 Grenoble and the <sup>§</sup>Ifremer, UMR6197, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France

**Background:** TET aminopeptidases are 12-subunit complexes present in the three domains of life and are involved in important biological functions.

**Results:** The TET assembling process has been characterized. The oligomerization triggers TET activity toward large polypeptidic substrates.

**Conclusion:** The assembling of TET is a controlled process and regulates its activity *in vivo*.

**Significance:** This work provides a new example of peptidase regulation driven by self-oligomerization.

Tetrahedral (TET) aminopeptidases are large polypeptide destruction machines present in prokaryotes and eukaryotes. Here, the rules governing their assembly into hollow 12-subunit tetrahedrons are addressed by using TET2 from *Pyrococcus horikoshii* (PhTET2) as a model. Point mutations allowed the capture of a stable, catalytically active precursor. Small angle x-ray scattering revealed that it is a dimer whose architecture in solution is identical to that determined by x-ray crystallography within the fully assembled TET particle. Small angle x-ray scattering also showed that the reconstituted PhTET2 dodecameric particle displayed the same quaternary structure and thermal stability as the wild-type complex. The PhTET2 assembly intermediates were characterized by analytical ultracentrifugation, native gel electrophoresis, and electron microscopy. They revealed that PhTET2 assembling is a highly ordered process in which hexamers represent the main intermediate. Peptide degradation assays demonstrated that oligomerization triggers the activity of the TET enzyme toward large polypeptidic substrates. Fractionation experiments in *Pyrococcus* and *Halobacterium* cells revealed that, *in vivo*, the dimeric precursor co-exists together with assembled TET complexes. Taken together, our observations explain the biological significance of TET oligomerization and suggest the existence of a functional regulation of the dimer-dodecamer equilibrium *in vivo*.

Efficient and controlled intracellular polypeptide breakdown is a primordial requirement that controls many cellular processes (1). Aminopeptidases constitute a group of enzymes of

critical importance to intracellular regulatory networks. In addition to their key role in energy and amino acid metabolism, they contribute in a crucial manner to the protein degradation pathways by trimming the peptides released by ATP-dependent proteases such as the proteasome (2–5). Moreover, aminopeptidases remove post-translationally the N-terminal amino acid from precursor proteins thus directing their maturation and cellular localization, as well as controlling their half-lives (6). These enzymes equally assume important roles in specific biological processes involving peptide signaling, such as the production of the major histocompatibility complex ligands (7). Therefore, altered intracellular aminopeptidase activities have been associated with a variety of pathologies, including aging, cataracts, cystic fibrosis, angiogenesis, and cancers (5, 8–10).

Numerous intracellular energy-independent peptidases co-exist in the cytosol, but only a few of them share the capacity for self-assembly as large homo-multimeric complexes. These include bleomycin hydrolase, leucine aminopeptidase, DppA, tripeptidyl peptidase II, Tricorn protease, protease1, pab87, and the TET<sup>4</sup> aminopeptidase (11–18). All these systems confine the peptidase activity to inner cavities, accessible exclusively to unfolded polypeptides. Although most of them adopt a barrel-shaped architecture in which the active sites are lined up alongside a single central channel, the TET aminopeptidases form unique dodecameric edifices with a typical tetrahedral shape (18). The TET particle interior is accessible via the openings situated on each facet of the tetrahedron. The internal organization of TET peptidases revealed a highly self-compartmentalized system comprising a crossing network of four access channels extended by vast catalytic chambers in which three active sites are arranged in a circular fashion (19, 20). This spatial arrangement is structurally different from all other self-

<sup>\*</sup> This work was supported in part by the Agence Nationale de la Recherche Grants "MacroTET"-BLAN-07-3 204002 and "Archelyse" ANR-12-BSV8-0019-01 and the GRD Ecchis.

<sup>[5]</sup> This article contains supplemental Fig. 1.

<sup>1</sup> Supported by a Ph.D. scholarship from the French Ministry for Research and Technology.

<sup>2</sup> Supported by a French National Research Agency postdoctoral fellowship.

<sup>3</sup> To whom correspondence should be addressed: Institut de Biologie Structurale, 41 Rue J. Horowitz, F-38027 Grenoble Cedex 1, France. Tel.: 0033-4-38-78-95-69; Fax: 0033-4-38-78-95-69; E-mail: franzetti@ibs.fr.

<sup>4</sup> The abbreviations used are: TET, tetrahedral aminopeptidase; SAXS, small angle x-ray scattering; AUC, analytical ultracentrifugation; PDB, Protein Data Bank; pNA, p-nitroanilide; TAPS, 3-[(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]-1-propanesulfonic acid.



compartmentalized protease complexes of known three-dimensional structure.

TET cytosolic enzymes belong to the M42 or M18 metallopeptidase families in the clan MH according to the MEROPS classification system (21). The typical TET dodecahedral quaternary structure was initially described in archaea (18, 22, 23). It was also found in bacteria (24), and recently, the crystallographic structure of bovine and human tetrahedral aspartyl-aminopeptidases have revealed that the TET complexes are also present in eukaryotic cells (25, 26). The quaternary structure of archaeal, bacterial, and eukaryal TET assemblies is highly conserved. Their tertiary structures show that they all exhibit a two-domain architecture consisting of a catalytic and a dimerization domain (19, 24, 25).

The high evolutionary conservation of TET peptidases in the three kingdoms of life suggests that they perform important biological functions. They were found to process polypeptides up to 27 amino acids in length in the absence of ATP (18, 27). Three different versions of TET complexes co-exist in the hyperthermophilic archaeon *Pyrococcus horikoshii* as follows: PhTET1, PhTET2, and PhTET3. These complexes are built up in an identical fashion and have comparable dimensions. PhTET2 can be defined as a leucyl-aminopeptidase that displays a preference for neutral and aliphatic substrates; PhTET3 is a lysyl-aminopeptidase that hydrolyzes preferentially basic residues, and PhTET1 is a glutamyl-aminopeptidase that shows high specificity toward acidic residues (19, 27, 28). The comparison of the surface electrostatic potential features of the proteolytic chambers and of the structures of the active site pockets suggest a mechanism of substrate (N-terminal amino acid) discrimination based on the PhTET internal surface electrostatic potential features (19, 28).

The particular substrate specificities of the three TET variants in *P. horikoshii* suggest that they form a complementary set of enzymes (28). Because of their cooperative action, the archaeal TET peptidases can be designated as a "peptidasome" involved in the destruction of a vast variety of polypeptides. It has been suggested that, in peptide-fermenting organisms, the TET system plays an important role in the energy metabolism (19) or in the intracellular protein degradation by hydrolyzing the peptides produced by the proteasome endopeptidase activity (23). In addition, the TET peptidases could play more specific physiological roles as they can cleave physiologically relevant peptides. This hypothesis is supported by recent work on the eukaryotic tetrahedral aspartyl aminopeptidase that has been proposed to be a key player in the central nervous system, in particular by regulating the ocular and renin system (25).

The TET enzymes are co-catalytic metallopeptidases typically binding, by means of five amino acid ligands, two atoms of zinc or cobalt per monomer. The catalytic mechanism also implies a glutamate and an aspartate residue. Cobalt ions have a clear stimulatory effect on the amidolytic activity of PhTETs, and the co-catalytic metals have been found to be important to maintain the PhTET oligomerization state (19, 27–29). Unlike other self-compartmentalized peptidases, TETs are not processive enzymes that imply the detachment of the peptide moiety from the active site once the N-terminal residue has been cleaved (19, 27, 28, 30). The mechanism of TET hydrolysis is

extremely similar to the one of secreted monomeric aminopeptidase such as *Vibrio* aminopeptidase Ap1; it is the charge properties and the dimensions of the catalytic pocket of each monomer that trigger the specificity of the enzyme toward the N-terminal amino acid from the peptide chain (31). Thus, in the case of TET peptidases, the biological significance for oligomerization and active site self-compartmentalization is not clear. To address this question, a site-directed mutagenesis strategy was used to slow down the natural oligomerization process of the PhTET2 complex. The structural properties of the purified PhTET2 dimer and of various oligomeric form intermediates were characterized by combining small angle x-ray scattering (SAXS), native gel electrophoresis, analytical ultracentrifugation (AUC), and electron microscopy. This allowed the dissection of the TET assembling pathway. The relationship between the aminopeptidase activity and its multimeric structure was also assessed by functional assays. Finally, density gradient fractionations and immunodetection experiments performed with *Halobacterium* and *Pyrococcus* cell extracts suggested the existence of a regulatory mechanism controlling the TET oligomerization state *in vivo*.

## EXPERIMENTAL PROCEDURES

***P. horikoshii* and *Halobacterium salinarum* Cell Cultivation**—*P. horikoshii* was grown on SME YP medium, pH 7 (32), supplemented with PIPES (20 mM) and elemental sulfur (1 g/liter). The cultivation was performed overnight at 90 °C in 1-liter serum vials containing 500 ml of medium under anaerobic conditions (N<sub>2</sub> gas phase). Cells were harvested by centrifugation, and cell pellets were immediately stored at room temperature in water/isopropyl alcohol (50:50 v/v) until utilization. *H. salinarum* NRC1 was grown in hypersaline medium and processed as described (33). Cells were harvested at the end of mid-log phase.

***PhTET2* Cloning, Expression, and Mutagenesis**—Wild-type and mutant (R217S, R220S, F224S, H248S, and I292A) *PhTet2* genes were generated from synthetic DNA fragments optimized for codon usage in *Escherichia coli* and cloned in the overexpression plasmid pET41c by GeneCust Europe. The resulting constructs were transformed in the *E. coli* strain BL21 (DE3) for the recombinant expression of wild-type and mutated PhTET2 proteins as described in Durá *et al.* (27).

**Protein Purification**—Total protein extracts from *E. coli* cells expressing the various recombinant PhTET2 proteins were prepared as described in Durá *et al.* (27). After heating at 85 °C for 15 min to eliminate most mesophilic proteins from the *E. coli* host, the lysates were clarified by centrifugation at 17,400 × *g* for 1 h, and the supernatant was loaded onto a 6-ml Resource Q column (GE Healthcare) equilibrated in 100 mM NaCl, 20 mM Tris-HCl, pH 7.5. After washing with 3 column volumes, bound proteins were eluted with a linear salt gradient (0.1–0.35 M NaCl in 20 mM Tris-HCl, pH 7.5). In the case of the PhTET2 pentamutant, two well separated elution peaks (called A and B) were observed. The fractions of each peak were combined and concentrated to 5 mg/ml using a centrifugal filter unit (Millipore) with a molecular mass cutoff of 10 kDa. For further purification and equilibrium shift assays, the proteins were loaded onto a Superose 6 size exclusion column (GE

## Assembling TET Peptidases

Healthcare) equilibrated in the desired buffer. To study salt (from 20 to 300 mM NaCl) and pH (7–9) effects, 50 mM HEPES and TAPS buffers were used. The purified proteins were kept at 4 °C after the size exclusion step. Aliquots were analyzed on native gel electrophoresis as described below to observe the evolution of the equilibrium with time. Native gel electrophoresis in combination with analytical size exclusion chromatography experiments (see below) were systematically performed to determine optimal buffer conditions for AUC, SAXS, and activity measurements and to monitor the oligomerization state of the PhTET2 protein samples before and after experiments.

**Native Gel Electrophoresis Experiments**—Native-PAGE experiments were carried out to accurately identify the different PhTET2 subspecies generated upon incubation at different times and buffer conditions. Protein samples were mixed with 1 volume of loading buffer (20 mM Tris-HCl, 62.5% glycerol, pH 6.9) just before analysis. Polyacrylamide gels containing 8% acrylamide/bisacrylamide, 20 mM NaCl, 50 mM Tris-HCl, pH 8.8, were run at 4 °C for 90 min in a Tris/glycine buffer (25 mM Tris-HCl, 192 mM glycine, pH 7.5). The protein bands were visualized by Coomassie Brilliant Blue staining. The ratios between the different PhTET2 species that were detected on native gel electrophoresis were found to be in excellent agreement with those determined on the same samples and in the same buffer by size exclusion experiments on a Superose 6 column.

**AUC**—Sedimentation velocity experiments were performed at 20 °C on different samples of the purified mutated PhTET2 protein with an Optima XL-1<sup>®</sup> analytical ultracentrifuge (Beckman) at 42,000 rpm with a 50-Ti eight-hole rotor (Beckman Instruments). The analyzed samples were prepared in 50 mM HEPES, 300 mM NaCl, pH 7.5 buffer, at a protein concentration of 0.4 mg/ml as described above. Two-channel centerpieces with an optical path of 12 mm were used, and all experiments were performed using sapphire windows. Scans were recorded at 280 nm with radial spacing of 0.005 cm. The program Sednterp was used to estimate the partial specific volume from amino acid composition as well as the density  $\rho$  and viscosity  $\eta$ . Data were analyzed with the program Sedfit (34) using a continuous  $c(s)$  distribution model.

**SAXS**—All experiments were performed on the beamline BM16 at the European Synchrotron Radiation Facility (Grenoble, France). Scattering curves of the samples were measured at 80 °C. The scattering patterns were corrected for background scattering and the geometry of the experimental arrangement. Three samples were measured as follows: the mutated dodecamer of PhTET2 (PhTET2-12s), purified to homogeneity just before the experiment and concentrated up to 2.9 mg/ml in 50 mM HEPES, 20 mM NaCl, pH 7.5; the mutated dimer of PhTET2 (PhTET2-2s), purified to homogeneity just before the experiment and concentrated up to 3 mg/ml in 50 mM TAPS, 20 mM NaCl, pH 9; and the wild-type dodecamer of PhTET2 (PhTET2 WT), purified to homogeneity just before the experiment and concentrated up to 3.5 mg/ml in 50 mM HEPES, 20 mM NaCl, pH 7.5. The *ab initio* envelope was generated using the program GASBOR (35). Fifteen models were averaged and filtered with the help of the program

DAMAVAR (36). The resulting bead model was transformed into a volumetric map using the program pdb2vol from the Situs program package (37) and the crystallographic model was placed into this map by rigid body docking with the program Colores (38). The theoretical SAXS curves were calculated with the program CRY SOL (39) in default mode. The missing fragments were modeled (internal loop, 120–132, and the N-terminal 1–5) in the PhTET2 12s structure and were added using PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC, and Coot (40).

**Electron Microscopy**—The dimeric fraction of PhTET2 was incubated in 50 mM HEPES, 150 mM NaCl, pH 7.5, after the size exclusion column step of the purification. The reoligomerization was followed by native gels and gel filtration. The intermediate fractions were isolated from the dimer and dodecamer of PhTET2 and pooled in a “low molecular weight intermediate” sample and a “high molecular weight intermediate” sample.

4  $\mu$ l of the PhTET2 sample ( $\sim$ 0.1 mg/ml) were loaded between the mica-carbon interface as described in Franzetti *et al.* (18). The sample was stained using 2% sodium silicotungstate, pH 7.5, and air-dried. Images were taken under low dose conditions in a CM12 Philips electron microscope working at 120 kV and with a nominal magnification of 40,000 using an Orius SC1000 CCD camera.

The different PDB files generated by molecular modelization were loaded into SPIDER (41) and filtered to 20-Å three-dimensional structures. Those structures were reprojected in all directions with a 10° spacing, and the reprojections were compared by cross-correlation with raw negative staining images selected from the CCD frames. For each structure, the best cross-correlating images were carefully examined by eye. The best matching and unambiguous images and the corresponding three-dimensional model orientation were selected and included in Fig. 5. Special care was taken to include only model-specific representative views (reprojections and views that are not overlapping with other model reprojection views).

**Activity on Monoacyl-pNA Peptides**—PhTET2 hydrolytic activity on different synthetic chromogenic compounds was determined using aminoacyl-pNAs as described (27). Four replicates and four enzyme blanks were assayed for each experimental point. Leu-pNA was used for half-life calculations as described (27). Half-lives were calculated from a first-order exponential decay fit to the experimental points.

**Activity on Polypeptides**—PhTET2 hydrolytic activity on synthetic peptides was determined through a procedure inspired from Frotin *et al.* (42). Different commercial polypeptides containing a poorly hydrolyzable residue in position P<sub>1</sub>' were assayed. Reactions were initiated by addition of the enzyme (0.1 mg/ml) to 140  $\mu$ l of a pre-warmed reaction mixture containing 0.1 mg/ml *o*-dianisidine (Sigma), 3 units of horseradish peroxidase (Sigma), 0.5 units of L-amino acid oxidase (Sigma), and 0.1–6 mM peptide in 50 mM PIPES, 150 mM KCl, 1 mM CoSO<sub>4</sub>, pH 7.5. The hydrolytic activity was assayed by monitoring the absorbance of oxidized *o*-dianisidine at 440 nm, coupling the PhTET2 activity to both L-amino acid oxidase and peroxidase activity, according to the reaction sequence described by Frotin *et al.* (42). Assays were performed in 1-mm-thick quartz cuvettes. Absorbance evolution was measured using a Beckman

spectrophotometer DU 7400 equipped with a thermostat and a 6-position sample changer. The hydrolytic activity of PhTET2 was measured at 40 °C. To calculate the concentration of oxidized *o*-dianisidine in solution, the molar extinction coefficient used was  $\epsilon = 10,580 \text{ M}^{-1} \text{ cm}^{-1}$ . Catalytic constants  $K_m$  and  $k_{\text{cat}}$  were estimated using the enzymology tools of SigmaPlot version 11.0 from Systat Software, Inc. For the same mass concentration, the molecular concentration of dimer is six times higher than the one of the dodecamer. The difference between the mass and the molar concentrations has been taken into account for the estimation of the kinetic parameters. Four replicates and four enzyme blanks were assayed for each experimental point.

**Sucrose Density Gradient Fractionations and Protein Immunodetection Experiments**—*P. horikoshii* cell pellets (0.2 g) were suspended in 1 ml of lysis buffer (150 mM KCl, 50 mM Tris-HCl, 40 mM MgCl<sub>2</sub>, DNase I grade I (Roche Applied Science, 0.05 mg/ml), pH 7.6). The disruption of the cells was achieved by sonication (10 pulses of 20 s at medium power with a Branson Sonifier 50). The crude extracts were clarified by centrifugation at  $30,000 \times g$  for 30 min at 4 °C (Ultracentrifuge Beckmann Optima-TL 100), and 250  $\mu\text{l}$  of the post-membrane protein cytoplasmic extracts (S30) was gently loaded onto 5–25% continuous sucrose density gradients made in 10 ml of Beckman ultraclear centrifuge tubes. The sucrose density gradients were centrifuged at  $210,000 \times g$  during 20 h (Beckman L-80 ultracentrifuge, SW-41 Ti rotor). *H. salinarum* cellular extracts and sucrose density gradient experiments (5–20%) were performed in native hypersaline conditions as described in Chamieh *et al.* (33).

Purified dodecameric and dimeric forms of the mutated PhTET2 protein were run simultaneously on separate tubes. The gradients were fractionated into 750- $\mu\text{l}$  aliquots. In *P. horikoshii* S30 fractionation experiments, the proteins were precipitated from the fractions during 30 min at –20 °C with trichloroacetic acid (15% final concentration). After centrifugation at  $16,000 \times g$  for 30 min at 4 °C, the pellets were washed with glacial acetone, and the samples were immediately centrifuged for 15 min at  $16,000 \times g$  at 4 °C. The pellets were air-dried for 10 min before resuspension in ice-cold 20 mM Tris base. One volume of SDS-PAGE loading buffer (50 mM Tris-HCl, 8 M urea, 2 M thiourea, 100 mM DTT, 3% SDS, 0.1% bromophenol blue, 10% glycerol, pH 6.8) was then added to the samples. These were heated at 95 °C for 5 min, and their proteins were separated by SDS-PAGE using 12% gels. Staining was done with Coomassie Brilliant Blue or the proteins were transferred onto Hybond-P PVDF-membranes, and immunoreactive bands were visualized by chemiluminescence as detailed by the supplier (ECL detection kit; GE Healthcare). The signal was detected using a Kodak Image station 4000 mm. Specific PhTET2 antibodies were raised against the following synthetic peptide chosen in the protein primary sequence DERDVDAT-VELMTKALENIHELKI. The anti-HsTET antibodies were raised against two peptides, TRGSQVRIETDDGPV and AHAGDRDSFGVSV. The antisera were used at 1:2500 and 1:10,000 dilution for PhTET2 and HsTET, respectively. The experiment was repeated at least three times with *P. horikoshii* and *H. salinarum* cell batches from different cultures.

## RESULTS

**Purification and Structural Characterization of a Dimeric PhTET2 Precursor**—For this study, the PhTET2 complex from *P. horikoshii* was chosen. The protein exhibits high thermal stability, and compared with other types of TET complexes, it possesses broader substrate specificity. Additionally, it can be purified in large amounts without adding tags that may create artifacts when studying the oligomerization process. When recombinant PhTET2 proteins are expressed in *E. coli*, only the dodecameric complex was detected in the soluble post heat-shock fractions. Moreover, once assembled, the TET particles are extremely robust *in vitro* and can only be broken down into smaller oligomeric forms by using extreme pH conditions and EDTA treatments (29). These harsh physicochemical treatments do not allow the functional characterization of the low molecular weight species; moreover, the process is only partially reversible and generates aggregates. We have therefore used a site-directed mutagenesis approach to study the TET assembling process and the influence of the enzyme's oligomerization state on its catalytic and thermal stability properties. The examination of the subunit interfaces in PhTET1, PhTET2, and PhTET3 according to their crystallographic structures, as well as the structural study of a 24-subunit PhTET1 complex, suggested that the dimer is the building block of the PhTET edifices (19, 28). Point mutations were therefore designed to weaken the interactions at the interfaces between the dimers within the dodecamer. Some of them were changed to serine instead of alanine to completely eliminate the possible formation of any hydrophobic interaction of the side chain methyl group of the alanine and also to make the interface more polar. However, neither single nor double mutations could appreciably alter the rate of TET dodecamer formation in the post heat-shocked supernatant of the transformed *E. coli* cells, and only after a combination of five mutations (R217S, R220S, F224S, H248S, and I292A) was a significant slowdown of the oligomerization process achieved. The five mutations are mostly located in the prominent intersubunit interaction regions located at the apices of the tetrahedron on residues participating in hydrophobic clusters but also in polar contacts (for details see supplemental Fig. 1).

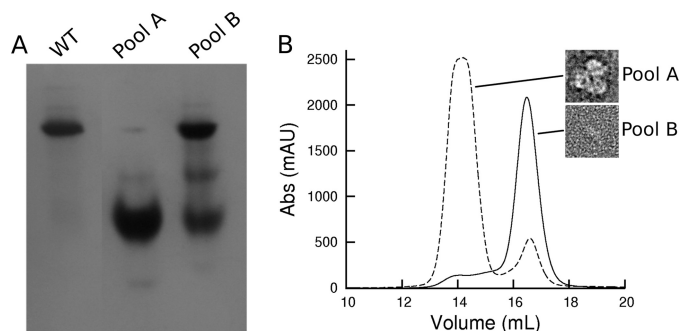
When the post-heat shock protein extracts of *E. coli* cells expressing the PhTET2 pentamutant were resolved on a Resource Q ion exchange column, two well separated elution peaks were obtained, whereas the expression of the wild-type PhTET2 protein produces only one peak. The corresponding fractions were pooled (pool A and pool B). Native gel electrophoresis analysis revealed that pool A contained a homogeneous population of low molecular weight oligomers, and pool B contained predominantly large oligomers that migrated as the wild-type TET dodecamer (Fig. 1A). The two oligomeric forms of the mutated PhTET2 proteins were further purified by size exclusion chromatography (Fig. 1B). The A pool contained PhTET2 oligomers with an apparent molecular mass of <100 kDa, whereas the B pool principally contained a large complex. According to the molecular weight calibration of the size exclusion column and to negative staining electron microscopy images, the low molecular weight form would correspond to a



## Assembling TET Peptidases

free TET dimer (78 kDa), although the larger complex was a dodecamer (468 kDa), as clearly shown by the negative staining electron microscopy pictures that revealed its tetrahedral shape, identical to that of the wild-type TET (Fig. 1B).

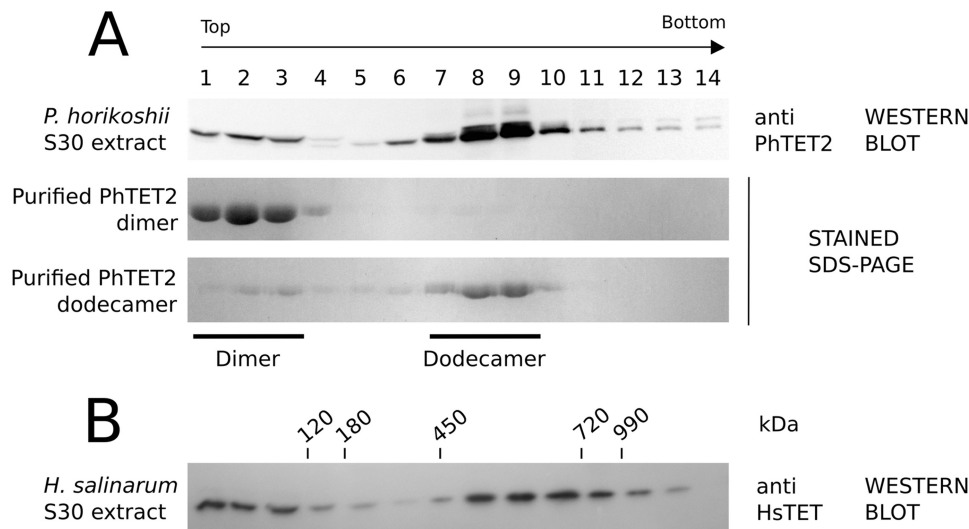
**TET Low Molecular Weight Complexes Co-exist with the Assembled Dodecamers in Vivo**—To gain insight into the oligomerization state of the PhTET2 aminopeptidase complex *in vivo*, post-membrane protein cytoplasmic extracts (S30) from *P. horikoshii* cells were resolved through continuous sucrose



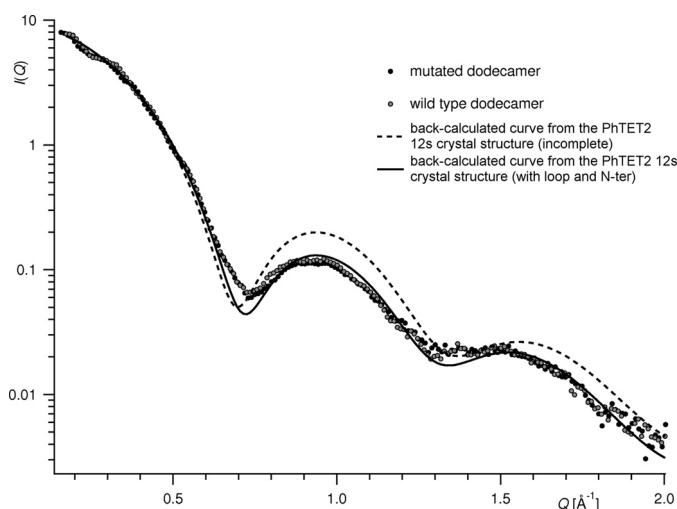
**FIGURE 1. Purification of a PhTET2 dimeric complex.** A, native gel electrophoresis analysis of the PhTET2 protein that was obtained after expression of the recombinant wild-type and the pentamutant (R217S, R220S, F224S, H248S, and I292A) proteins in *E. coli*, heat shock clarification and separation by Resource Q ion exchange chromatography. WT is the purified wild-type PhTET2 dodecamer; Pool A corresponds to a low molecular weight complex, although the major constituent of Pool B is a high molecular weight complex that migrates as the wild-type PhTET2 dodecamer. B, Superose 6 column chromatography purification of dimeric and dodecameric mutant PhTET2 complexes. The  $A_{280}$  absorption profiles are presented together with negative staining electron microscopy images taken from aliquots of the pooled peak fractions. The elution volume and the shape of the purified complexes from pools A and B correspond to dimers and TET dodecamers, respectively. This information was further confirmed by AUC (Fig. 5) and SAXS (Figs. 3 and 4).

density gradients. Fractions were collected, and the PhTET2 proteins were immunodetected in each fraction by using a specific antibody (Fig. 2A). Two major populations of native PhTET2 proteins complexes were found to accumulate in the cellular extracts, a high and a low molecular weight species. To assess their respective oligomerization state, purified pentamutant PhTET2 dodecamer and dimers were analyzed on separate gradients, in the same experimental conditions. The dimeric form was obtained from a PhTET2 mutant that is described in this study. It showed that the high molecular weight complex detected in *P. horikoshii* sedimented as the purified recombinant dodecamer, although the low molecular weight species would correspond to a PhTET2 dimer (Fig. 2A). This indicates that the five mutations introduced in the PhTET2 interfaces indeed lead to the stabilization of the low molecular weight form observed *in vivo* (Fig. 2A). This result is interesting because, when wild-type PhTET2 proteins are expressed in *E. coli*, only the dodecameric complex is detected in the soluble post heat shock extract.

To confirm the presence of unassembled TET complexes in the cytosol of archaea, we performed similar cell fractionation and immunodetection experiments in *H. salinarum*, an extreme halophilic organism that accumulates multimolar salt concentration in its cytosol (43). For this reason, all experiments on *H. salinarum* samples, including sucrose gradients sedimentation, were performed in physiological hypersaline conditions (3.4 M KCl) as described in Chamieh *et al.* (33). The results are shown in Fig. 2B. As in the *P. horikoshii* extracts, a low molecular weight species of TET with an apparent molecular mass of <100 kDa is present together with the fully assembled TET particle. We therefore concluded that, *in vivo*, a significant population of low molecular weight TET



**FIGURE 2. Determination of TET oligomerization states in *P. horikoshii* and *H. salinarum* cells.** A, immunodetection of the native TET2 complexes from *P. horikoshii* (Ph). Post-membrane extracts (S30) were loaded on a 5–25% sucrose density gradient and resolved by ultracentrifugation. The fraction numbers are indicated on top, and the arrow represents the sedimentation orientation. The proteins from each fraction were separated on a 12% SDS-polyacrylamide gel, and the presence of the PhTET2 protein was revealed by Western blot by using specific antibodies. The bottom panels show Coomassie Blue staining of 12% SDS-polyacrylamide gels. In this experiment, purified dimeric and dodecameric PhTET2 complexes were resolved on a sucrose gradient by using the same protocol as for the S30 extracts. These complexes were obtained from a mutant protein in which amino acids have been changed to slow down the particle assembling process (see Fig. 1). The positions of the dimer and the dodecamer in the gradients are indicated and correspond to the immunodetected PhTET2 complexes in the cell extracts. B, immunodetection of the native TET complexes from the extreme halophilic strain *H. salinarum*. S30 extracts were fractionated in native hypersaline conditions on a 5–20% sucrose gradient. The gradient was calibrated with halophilic protein complexes of known molecular sizes (indicated above) as described in Chamieh *et al.* (33). A low molecular weight TET complex co-exists with the assembled TET dodecamer in the cytosols of *P. horikoshii* and *H. salinarum*.



**FIGURE 3. SAXS study of the PhTET2 mutant and wild-type dodecamers.** SAXS curves from both particles superpose very well over the whole angular range studied, indicating that the quaternary arrangements of both particles are identical in solution. The back-calculated SAXS curves (using CRYSOLOG) from the 12-s PhTET2 crystal structure without and with missing fragments (internal loop and N terminus) are also shown.

complexes, presumably dimers, co-exists with the assembled TET dodecamers.

**Quaternary Structure and Thermal Stability Comparisons between the Wild-type and the Pentamutant PhTET2 Dodecamers**—The presence of low molecular weight forms of the TET complex *in vivo* prompted us to study their structural and biochemical properties as well as the TET assembling mechanism. The structural and biophysical properties of the mutant dodecameric particle were therefore studied to determine the potential effects of the mutations on its final quaternary edifice. To this end, SAXS curves were recorded on the mutant dodecamer in solution and compared with those of the PhTET2 wild-type protein. Both sets superposed nicely with the SAXS data from the wild-type TET, even when the experiments were performed at 80 °C, thus showing that the quaternary structure of the mutant is identical and equally stable as one of the wild-type TETs at the extreme physiological temperatures in which *P. horikoshii* thrives (Fig. 3). The agreement of the theoretical SAXS curve (calculated with CRYSOLOG) with the experimental data is good, in particular when the missing loop and the N terminus (not visible by x-ray crystallography) were modeled. The remaining discrepancy at the side minima could be explained by the conformational flexibility of these two fragments that produce a smearing of side minima in SAXS curves of globular objects (Lindner and Zemb (44)).

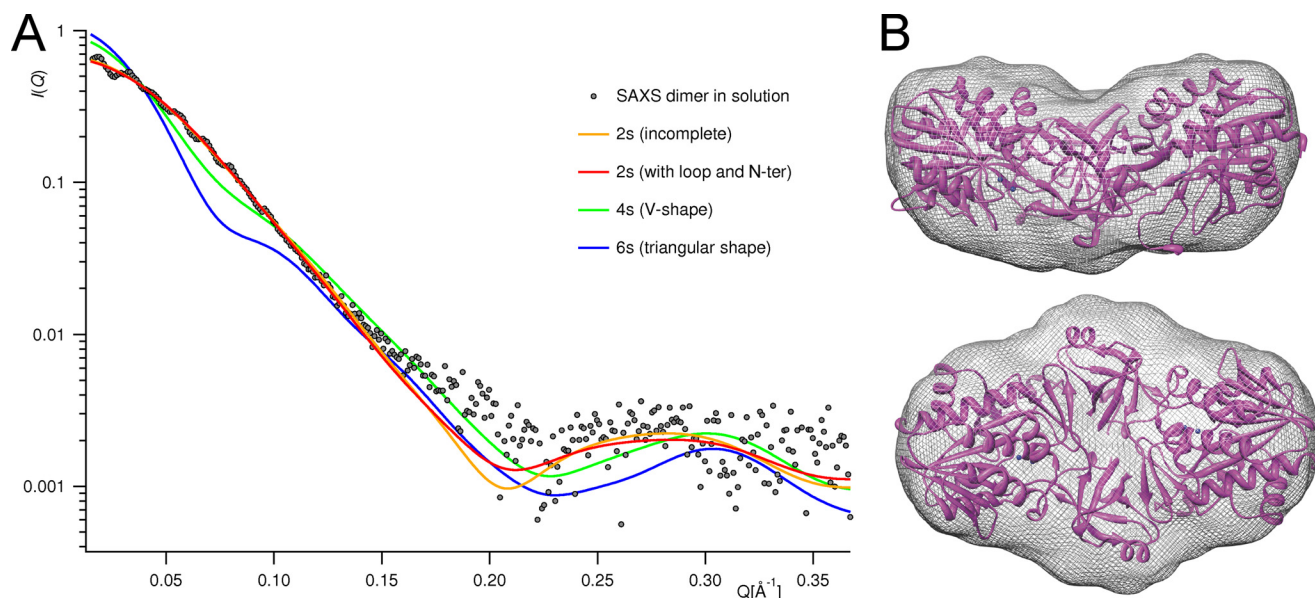
These experiments show that WT and mutant PhTET2 particles are therefore extremely stable once they are formed, and the deleted interactions located at the dimer-dimer interfaces do not contribute to the stability of the dodecameric structure at high temperatures.

**SAXS Structure of the Free TET Dimer**—Because the dimer is the basic building unit in the dodecameric TET particle (19), it is important to characterize its structure further to understand the mechanisms underlying the TET oligomerization process and to address the question of the physiological significance of unassembled dimers *in vivo*. Although crystallographic studies

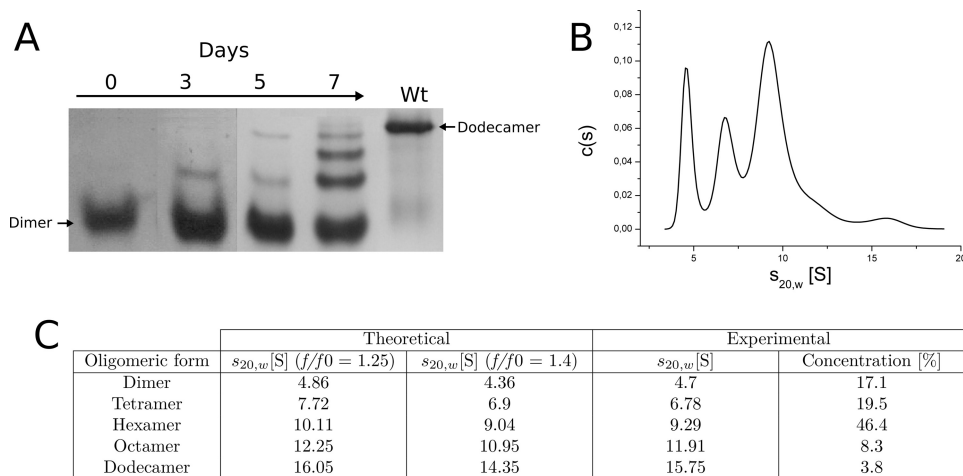
of the mutated free dimer are not possible due to its self-assembling properties, its oligomerization kinetics are slow enough to allow SAXS studies on monodisperse dimeric samples, even at 80 °C. The SAXS experimental curve obtained from the free dimer superposed nicely, up to a scattering vector  $Q = 0.15 \text{ \AA}^{-1}$ , with that calculated (with CRYSOLOG) from the crystallographic dimer structure within the wild-type dodecameric assembly (PDB entry 1Y0R) (Fig. 3), in particular when the missing internal loop and the N terminus were added. The minor mismatch around  $Q = 0.20 \text{ \AA}^{-1}$  can probably be explained by a conformational flexibility of the loop and N terminus as in the case of the dodecameric structure in Fig. 3. Moreover, the crystallographic structure of the dimer can be embedded very nicely into the *ab initio* envelope (Fig. 4B). A dimer of dimers (tetrameric V-shaped structure) and a trimer of dimers (hexameric triangularly shaped structure) did not fit the experimental SAXS data. We therefore conclude that the structure of the free dimer in solution is very similar to that of the wild-type dimer within the dodecamer. Thus, no major conformational changes of the TET dimer precursor seem to be necessary to constitute the 12-subunit tetrahedral particle.

**Biophysical Characterization of the PhTET2 Oligomerization Process**—The results described above showed that the mutated PhTET2 is fully able to self-assemble into a *bona fide* TET machine that maintains the same quaternary structure and thermal stability properties as the wild-type TET. Consequently, the pentamutant represents a model to study the oligomerization process. To identify the parameters controlling the equilibrium between the different oligomeric forms, samples were incubated in different buffers after the size exclusion purification step (see “Experimental Procedures”). At physiological pH, the formation of dodecamers from the dimer can be stimulated by a slight increase in salt concentration, although the dimer is stabilized at high pH. Native gel electrophoresis analysis performed at different times after incubation of the purified dimer revealed well defined bands corresponding to intermediate oligomeric states between the dimer and the dodecamer (Fig. 5A). The different oligomeric species were characterized by AUC (Fig. 5, B and C). Native gels electrophoresis and AUC analyses showed the same number of intermediates. Sedimentation profiles displayed multiple different peaks at 4.70, 6.78, 9.29, 11.91, and 15.75 S, indicating the presence of five species in the sample. The same number of species was observed on native gel electrophoresis. Using the Svedberg equation (45), we calculated theoretical sedimentation coefficients for dimeric and dodecameric PhTET2 as well as for putative intermediates (Fig. 5C). Accordingly, the peaks were assigned to, respectively, dimer, tetramer, hexamer, octamer, and dodecamer of PhTET2.

The native gel analysis of the different assembling intermediates revealed that the tetrameric form is the first one to be detected. Interestingly, the accumulation of the octamer precedes the apparition of the hexamer; we could not detect the decamer, and no intermediate forms with odd subunit numbers were identified (Fig. 5C). These experiments demonstrated that the TET oligomerization process is a nonrandom stepwise process.



**FIGURE 4. Low resolution structure of the free PhTET2 dimer.** *A*, experimental SAXS curves for the mutated PhTET2 dimer in solution. The back-calculated SAXS curves (using CRYSOLOG) from the 2-s PhTET2 with and without missing fragments (internal loop and N terminus) as well as the back-calculated SAXS curves from a V-shaped tetramer and a triangular hexamer are also shown. The experimental SAXS curve superposes well with the back-calculated one from the 2-s particle (including the missing loop and the N terminus), indicating that the quaternary arrangements of both particles are very similar in solution. *B*, overlay of the final averaged *ab initio* shape reconstruction of the mutated PhTET2 dimer derived from SAXS experiments (gray envelope) with the structure of the PhTET2 dimer inside the TET dodecamer crystallographic model (PDB entry 1Y0R, purple).

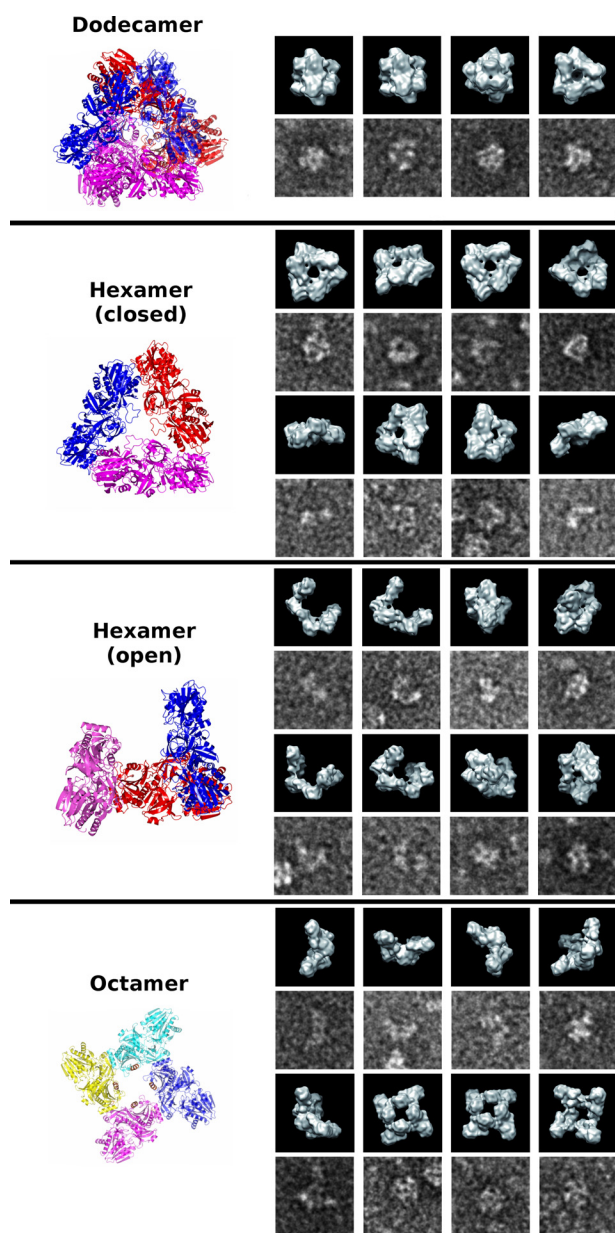


**FIGURE 5. Biophysical characterization of the PhTET2 assembling pathway.** *A*, native gel electrophoresis experiment revealing the existence of various intermediate TET oligomeric species after incubation of the purified PhTET2 dimer during different times, at 4 °C, in 50 mM HEPES, pH 7.5, 300 mM NaCl. *B*, analysis of a sample of purified mutant PhTET2 dimer by analytical ultracentrifugation 7 days after purification. The sample loading concentration was 0.4 mg/ml. Sedimentation coefficients are expressed in Svedberg units, 1 S = 10<sup>-13</sup> s. Three major species are detected at 4.70, 6.78, and 9.29 S, and two other peaks are visible around 11.91 and 15.75 S. *C*, theoretical sedimentation coefficients for putative PhTET2 assembly intermediates calculated for two different frictional ratios and experimental sedimentation coefficients and respective relative concentration for species detected in the sample analyzed by AUC. The experimental sedimentation coefficients were calculated with the program Sedfit using a continuous *c(s)* distribution model. Only intermediate forms with even subunit numbers were detected thus indicating that the TET oligomerization is a nonrandom controlled process. The tetrameric and hexameric forms are found to be the most abundant in the analyzed sample.

**Electron Microscopy and Structural Modeling of the Different PhTET2 Intermediates**—The PhTET2 assembling intermediates identified by AUC and native gels were separated from the dimer precursors and from the final TET dodecamer by using gel filtration chromatography. The column fractions were analyzed by electron microscopy. Three types of well structured edifices were identified in the samples as follows: a tricorner-shaped complex, an open chain, and a complex with a 4-fold symmetry. The sizes of these complexes were consistent with the hexameric and octameric intermediates complexes.

Smaller, less structured species that would correspond to the tetramers could also be seen. To propose structural models for the octameric and hexameric TET assembling intermediates, we generated the dimeric building blocks from the crystal structures of PhTET2 (PDB code 1Y0R). By using the oligomerization interfaces that were identified in the PhTET2 dodecamers, it was possible to generate a tricorner complex made of three dimers (Fig. 6). Alternatively, three dimers can also assemble as an open chain with a Z shape (Fig. 6). These two assemblies can be recognized in the dodecamer.





**FIGURE 6. Comparison between the different oligomeric models of PhTET2 and the negative staining electron microscopy images.** The atomic resolution structure of the different PhTET2 assembling intermediates plus the real dodecameric assembly are represented on the left. For each of these particles a 20-Å resolution three-dimensional structure has been calculated. Some typical views obtained by negative staining electron microscopy (lower row) and the three-dimensional corresponding isosurface view (top row) are shown for each category. The dimensions of the negative staining squared images are  $34 \times 34$  nm, and the three-dimensional models images are  $20 \times 20$  nm.

The existence of complexes with a 4-fold symmetry cannot be modeled from the 12-subunit dodecamer structure. We used instead the 24-subunit structure that has been described for PhTET1 (19). Indeed, this tetraicosameric edifice is not fashioned through the assembly of two dodecamers, but the octamer is one vertex of the 24-mer assembly. The octamer uses different intersubunit contacts than the dodecamer. In particular, the interface with the five proposed mutations is not involved in the octamer. As the root mean square deviations between the PhTET1 (PDB code 2CF4) and PhTET2 dimers

were 5.1 Å over 660 C-α atoms, we used the PhTET2 dimer for modeling octamer. Accordingly, we extracted an octameric structure from the PhTET1 24-subunit one and superimposed the PhTET2 dimer structure to produce a PhTET2 octamer model. The structures corresponding to the octameric and hexameric models are presented in Fig. 6. Low resolution envelopes were generated from the corresponding coordinates files for back projection using Spider (41). The experimental electron microscopy images were compared with the different views of the complexes (see “Experimental Procedures”) (Fig. 6). As a control, the same work was performed with the final TET particle to clearly distinguish the differences between the shapes and dimensions of the tricorn and the TET dodecamer, respectively. Model-specific images could be found for each orientation of the complexes. The shapes and dimensions of the single particles identified by electron microscopy are consistent with those arising from the modeling, the AUC, and the native gel experiments. These results indicate that the proposed octameric and hexameric (tricorn and open chain) structures correspond to the octameric and hexameric intermediates that were identified in the PhTET2 assembling pathway.

The octameric square complex is built upon different interfaces to the two hexameric forms and the final TET dodecamer. The same interactions can be used to generate the tetramer intermediates. The five mutations that were designed to slow down the TET assembling process do not impact the alternative contact area present in the tetramers and octamers. Accordingly, the mutations favor the accumulation of tetramers and octamers, although the efficient formation of hexamers is delayed. This explains why the tetramers and octamers preceded the apparition of the hexamers during the assembling process of the PhTET2 pentamutant. It is noteworthy that the formation of the TET dodecameric apices that possess a 3-fold symmetry cannot arise directly from the tetrameric or octameric structures. This indicates that the tetramers and octamers are the products of an alternative pathway induced by the mutations.

*Oligomerization Triggers PhTET2 Activity toward Large Polypeptidic Substrates*—To study the role of the oligomerization with respect to the functional state of the aminopeptidase, the hydrolytic activities of the mutated PhTET2 dimer and dodecamer were first measured by using short chromogenic aminoacyl compounds as described by Durá *et al.* (27). These experiments were performed on a time scale of several minutes after the size exclusion purification step. As a control, aliquots were taken from the reaction mixture and analyzed by size exclusion chromatography and native gel electrophoresis to be sure that the dimers and dodecamers had not formed other multimers during the experiments. The rate of Leu-pNA cleavage performed by the enzyme dimeric forms and its comparison with the one of the mutant dodecamers revealed that, at 40 and 80 °C, the activity of the two types of oligomers is nearly identical. The  $K_m$  and  $k_{cat}$  values were found to be very similar for the two types of TET assemblies (Table 1). This study shows that the dimeric form of TET is active and that the assembling of the protein into a 12-subunit tetrahedral superstructure does not modify significantly its amidolytic activity toward small substrates.



The aminopeptidase capability of the dimer raised the question of the biological significance of the TET supramolecular complex. We previously showed that TET peptidases display a significant hydrolytic activity toward the N-terminal residue of oligopeptides up to 27 amino acids in size (27). When the activities of the mutant PhTET2 dodecamer and dimer were tested against longer peptides such as AAA-*p*NA and AAAA-*p*NA, a significant lag phase was observed in the case of the dimer. This suggests that the size of the peptide is an important parameter for substrate hydrolysis when comparing dimers to the fully assembled TET complex. To obtain kinetic constants from long polypeptides (about 10 amino acids in length), substrates that contained a favorable N-terminal residue and an unfavorable residue in position P<sub>1</sub>' were used. In this manner, the quantity of released amino acids reflects only the hydrolysis of the first residue, allowing the calculation of the kinetic parameters. The enzymatic assay employed was inspired from that developed by

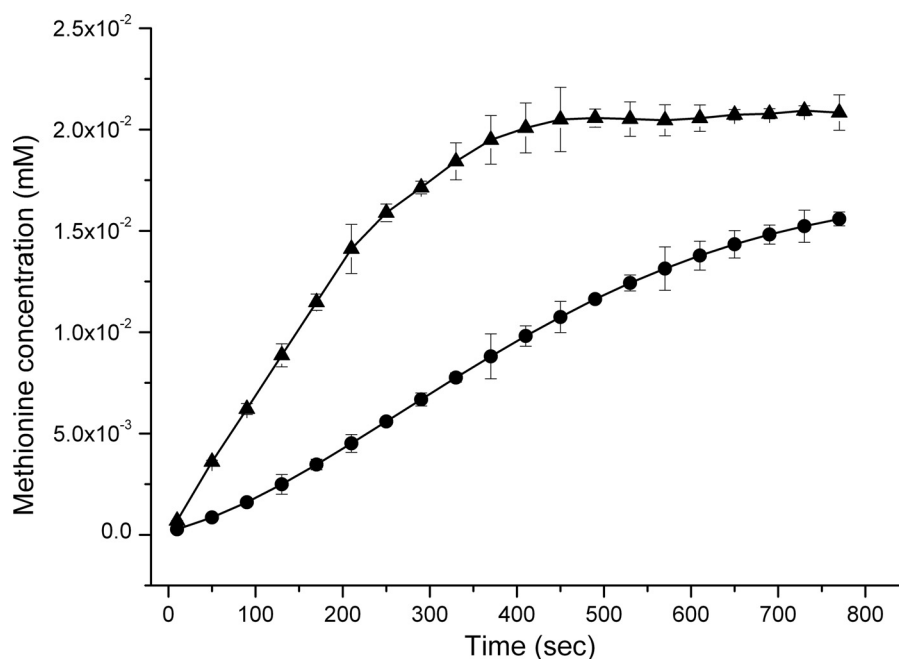
**TABLE 1**  
Kinetics constant of mutated PhTET2 dimer and dodecamer against a short (Leu-*p*NA) and an 11-residue (Met-Lys-bradykinin) peptide

Kinetics constant of mutated PhTET2 dimer and dodecamer against a short (Leu-*p*NA) and an 11-residue (Met-Lys-bradykinin) peptide. The  $k_{\text{cat}}$  is the catalytic efficiency of the enzyme; the  $K_m$  is the Michaelis constant of the enzyme reflecting the affinity of the enzyme for its substrate. The catalytic efficiency of the dimer on a long peptide is lower than the one of the dodecamer. The dodecameric form of PhTET2 is more efficient at processing long peptides than the dimeric form.

	Dodecamer	Dimer
Leu- <i>p</i> NA		
$k_{\text{cat}}$ (s <sup>-1</sup> )	511 ± 45.9	421.5 ± 33.6
$K_m$ (mM)	1 ± 0.1	1.1 ± 0.09
$k_{\text{cat}}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )	511 ± 97.09	383.2 ± 61.7
Met-Lys-bradykinin		
$k_{\text{cat}}$ (s <sup>-1</sup> )	144.5 ± 9.4	1.5 ± 0.13
$K_m$ (mM)	1.9 ± 0.114	2 ± 0.07
$k_{\text{cat}}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )	76.05 ± 9.5	0.55 ± 0.06

Frotin *et al.* (42). Again, the oligomeric state of the TET aminopeptidase was checked before and after the experiments. Two oligopeptides were tested, Met-Lys-bradykinin (MKRPPGF-SPFR) (Fig. 7) and MBL peptide <sup>176</sup>VDLTGNRLTY<sup>185</sup> (data not shown). In both cases, we observed that the dodecamer was more active than the dimer, and unlike for the short peptide experiment, an important difference in the reaction speed between the two PhTET2 oligomeric forms was observed. We measured the  $K_m$  and  $k_{\text{cat}}$  values of the dimer and dodecamer of the PhTET2 pentamutant for the Met-Lys-bradykinin peptide (Table 1). The  $K_m$  values of the dimer and the dodecamer were found to be identical. However, the  $k_{\text{cat}}$  value of the dimer was significantly lower than the one of the dodecamer. Thus, the TET reaction efficiency toward long peptides is compromised when the enzyme is not assembled as a dodecameric complex. These results demonstrate that oligomerization allows the TET peptidase to better process long polypeptides. Therefore, the dimeric and dodecameric forms that we detected *in vivo* in two archaeal strains (Fig. 2, A and B) may perform different physiological roles.

**PhTET2 Dimer Is Stable under High Temperature Conditions—**The TET dimer represents the minimal assembling unit in the TET oligomerization pathway (19), and this observation raises the question of the dimer stability under the extreme temperature conditions that prevail in the natural environment in which *Pyrococcus* cells thrive. The SAXS experiments described here have demonstrated that, *in vitro*, the dimer population does not show any sign of structural alteration at 80 °C, within 1 h (Fig. 4A). To further explore the thermal stability of the TET dimer, the residual aminopeptidase activity after different times of incubation at 80 °C was measured as described previously (27). Leu *p*NA was used as a substrate. Aliquots were



**FIGURE 7. Time course of hydrolysis of Met-Lys-bradykinin by the dodecameric and dimeric forms of the PhTET2 pentamutant.** Concentration evolution of oxidized *o*-dianisidine in the reaction mixture as a function of time is shown. For each free amino acid released in solution, one molecule of *o*-dianisidine is oxidized (see "Experimental Procedures"). Because PhTET2 displays very low activity toward basic residues such as Lys, the concentration evolution of oxidized *o*-dianisidine in solution reflects the activity of PhTET2 only on the full-length 11-amino acid substrate peptide. The corresponding kinetics constants are presented in Table 1.

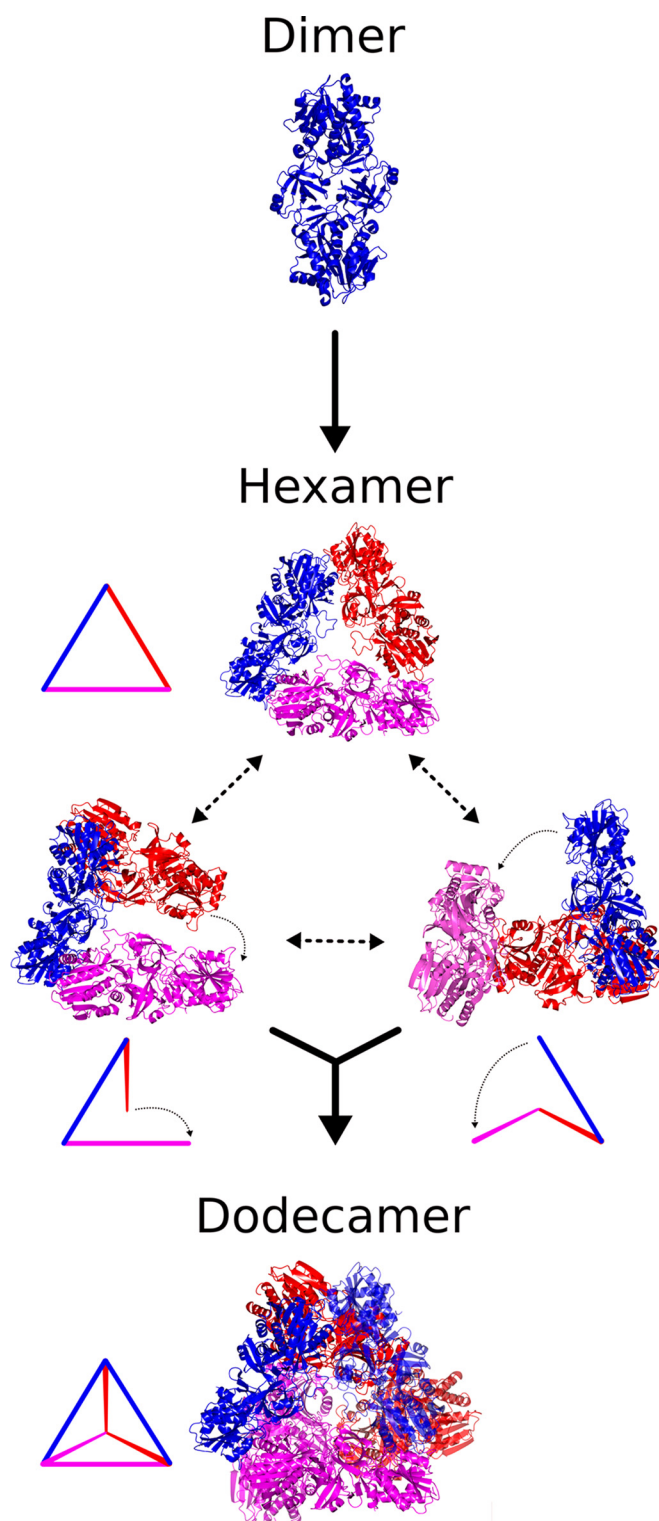
taken at different incubation times. Native gel and gel filtration experiments were performed to verify that no evolution of the protein oligomeric state had occurred during the incubation time. The calculated half-life values of the mutant dimeric and dodecameric forms were found to be the same as those of the wild-type PhTET2 determined earlier (about 10 h) (27). These findings indicated that the extreme thermal stabilization of the TET particles is not achieved through dodecamerization and that, under extreme physiological temperatures, free TET dimers, which are catalytically active against small peptides, could be accumulated in the *Pyrococcus* cells as precursors of the TET complex.

## DISCUSSION

A key question regarding self-compartmentalized peptidase complexes are the mechanisms by which such well organized edifices assemble. In this study, a site-directed mutagenesis strategy has been used to slow down the self-assembling process of TET, a large dodecameric aminopeptidase present in the three domains of life. The disruption of stabilizing interactions located in the interface area between the subunits at the apices of the tetrahedron had a significant effect on the kinetics of the TET assembling process. This way, a stable dimeric species could be purified that was fully able to self-assemble into active dodecamers. SAXS studies showed that the dodecameric particle obtained from the mutant dimer displayed the same stability and quaternary structure as the wild-type TET complex, even under extreme temperatures.

Native gel electrophoresis, AUC, and electron microscopy analyses proved that the mutant PhTET2 oligomerization does not proceed by an incremental addition of dimers but involves well defined intermediate species that are generated with time from the PhTET2 dimer precursor. Although tetramers and octamers were detected at the beginning of the oligomerization process, they are the products of an alternative pathway induced by the mutations. Based on SAXS analysis of the free PhTET2 dimer, electron microscopy study, and structural modeling from existing crystallographic structures of the PhTETs (19, 23), we propose that the natural pathway involves two hexameric intermediates, a tricorn and an open (Z-form) chain complex. Indeed, the tricorn represents a stable oligomeric form in solution. However, the Z-form is able to self-associate, leading to the extended interaction zone involving the catalytic domains of three subunits present in the highly stable biological dodecamers. Consequently, we propose that the open and closed hexamers are in equilibrium and represent the intermediates in the PhTET2 dodecamer assembling process (Fig. 8). The oligomerization interface that is present in the PhTET2 hexamers and dodecamers is well conserved in the three PhTETs (data not shown). This suggests that the oligomerization process and the associated functional activation that we described for PhTET2 may also be valid for PhTET1 and PhTET3.

Self-compartmentalization is a hallmark of cytosolic peptidases that degrade peptides in an unspecific manner (46). In the proteasome, tricorn protease, and bleomycin hydrolase, the sequestration of the catalytic sites from the bulk environment into chambers represents a peptide filtering system that is necessary to prevent unwanted damage on folded polypeptides within the cytosol (47–49). Oligomerization has also been



**FIGURE 8. Oligomerization model of the dodecameric TET particles based on the results obtained by the combination of site-directed mutagenesis, SAXS, AUC, electron microscopy, and structural analysis described here.** The dimer constitutes the building block and self-assembles into a closed hexamer, which involves the association of three dimers (red, magenta and blue). Two open "mirror" conformations of this complex are in equilibrium and finally associate to form the super-stable dodecameric particle.

shown to control different aspects of the peptidase functions. In the 20 S proteasome and its bacterial homologue ClpP, the priming of the peptidase enzymatic activity is coupled with sub-

unit association (50). In ATP-independent peptidases, such as protease 1 and Pab87, the active sites are formed at the subunit junctions (16, 17). In tripeptidyl peptidase II, a giant aminopeptidase found in eukaryotes, the activity increases in a nonlinear fashion with the oligomerization rate (11), and in the case of bovine lens leucyl aminopeptidase, it has been suggested that the activity depends on the stabilization of each monomer catalytic site by the structure of the oligomer (51). On the contrary, in  $\alpha$ -1 tryptase or kallikrein-related peptidases, self-oligomerization provides an inhibition of the enzyme activity (52, 53). With respect to these findings, it was surprising to find out that the dimeric form of the TET complex still carries out the same catalytic activity on small peptides as the 12-subunit complex. This indicates that, in the TET dimers, the catalytic sites and pockets are already positioned in proteolytically active conformations. The SAXS structure of the free dimer is consistent with these findings, because it shows that the relative positions of the two monomers of each dimer within the tetrahedral complex are already imposed by the interactions between dimerization domains. In the case of this small peptide, the calculated  $K_m$  values are similar, meaning that the substrate is recognized by the dimer and the dodecamer equally well. The similar calculated  $k_{cat}$  value reflects that the active site remains unchanged upon oligomerization. Thus, in the case of PhTET2, the oligomerization does not affect the active site and catalytic pocket configurations and has little effect on the recognition and trimming of N-terminal amino acids. However, when studying the catalytic activity of TET dimers and dodecamers as a function of substrate length, we found that the dodecamers are more efficient in hydrolyzing long polypeptide substrates as compared with the dimers. The kinetic parameters indicated that the dodecamer possesses a better efficiency than the dimer toward long substrates as follows: the  $k_{cat}$  value of the amidohydrolytic reaction, reflecting the catalytic efficiency of the system, is considerably reduced for the free dimer, although the  $K_m$  value, reflecting the affinity of the active site and the catalytic pocket for the N-terminal amino acid, remains the same between the dimer and the dodecamer. Thus, during the reaction, the same number of peptidase-substrate complexes are formed, but these complexes are more productive in the case of the dodecamer. In contrast, the peptidase-substrate complexes formed by the dimer are unproductive and are not true Michaelis-Menten complexes and therefore cannot be turned over. Therefore, in the case of the PhTET complex, the self-compartmentalization provides a way to enhance the enzyme efficiency when the substrate size increase. This represents another type of peptidase functional regulation driven by self-oligomerization. The interior of the TET peptidases consists of four polypeptide navigation channels that cross the particles from the entry pores situated on the facets of the tetrahedrons toward the catalytic chambers located within each apex (19, 28). A model for polypeptide processing was proposed based on the structural and enzymatic comparisons of the three TET enzymes from *P. horikoshii*. In this model, a series of mobile loops and electrostatic attractions/repulsions in the entry channels and in the catalytic chambers would orient the N terminus of the polypeptides toward the negatively charged active sites (19, 28). In unassembled TET dimers, the polypeptide naviga-

tion system is not present. This would explain why the TET dimer is disabled to break down long polypeptides.

*P. horikoshii* is a hyperthermophilic microorganism that grows optimally at 95 °C (54). In this work, it has also been shown that, *in vitro*, the free dimers are as stable at physiological extreme temperatures as the assembled TET particles, with half-lives of several hours at 80 °C. Different strategies are adopted by thermozymes to stabilize their native conformation at extreme temperatures (55). Among these, oligomerization has been proposed to represent an important determinant (56). In the case of the TET system, our results show that the dodecameric quaternary structure has little role in the high thermal stability of the enzyme. In fact, the monomer-monomer interfaces in the TET dimer consist of extended networks of ionic bonds that are likely to contribute to the high thermal stability of the enzyme as a free dimer. Therefore, dimers can be accumulated as stable precursors of the TET dodecamers *in vivo*.

There is little information available about the *in vivo* oligomeric states of large energy-independent peptidase complexes. We showed here that the PhTET2 dimer co-exists with the dodecamer in cellular extracts. A similar observation was also made under hypersaline conditions in the extreme halophilic archaeon *H. salinarum*. Unlike *Pyrococcus*, the *Halobacterium* genome contains only one copy of the TET peptidase. Thus, it is reasonable to propose that the dimer-dodecamer equilibrium is a hallmark for TET peptidases and that a specific regulation occurs *in vivo* to control the oligomerization state of TET. Because oligomerization affects PhTET2 activity with respect to the size of the polypeptidic substrates, a regulated oligomerization would therefore allow the TET particle to process a broad variety of peptides, from dipeptides to long polypeptides, in response to varying physiological demands; although the TET dimer would act preferentially on small peptides for energetic and anabolic purposes, its assembly into dodecamers and the concomitant formation of the polypeptide navigation system would trigger the intracellular aminopeptidase activity toward longer peptides such as those produced by the 20 S proteasome or toward peptides possessing specific biological activities. Thus, the control of the amount of assembled TET *in vivo* may represent an important regulatory step in protein degradation and in many specific biological functions based on polypeptide activity. Regarding that, *in vitro*, the TET oligomerization appears to be a rapid process, and given that the assembled TET particles are extremely robust, it is likely that the *in vivo* regulatory mechanisms involve the stabilization of the dimeric species. Studies are in progress in our laboratory to identify this mechanism.

**Acknowledgments**—We thank A. Le Roy and C. Ebel from the PSB/IBS platform for the assistance and access to the instrument of analytical ultracentrifugation. We thank D. Fenel and C. Moriscot from the PSB/IBS for the electron microscopy. We thank J. Le Bars for help in *P. horikoshii* cultivation. This work used the platforms of the Grenoble Instruct Center, UMS 3518 CNRS-CEA-UJF-EMBL.

## REFERENCES

1. Varshavsky, A. (2005) Regulated protein degradation. *Trends Biochem. Sci.* **30**, 283–286
2. Goldberg, A. L. (2003) Protein degradation and protection against mis-



- folded or damaged proteins. *Nature* **426**, 895–899
3. Saric, T., Graef, C. I., and Goldberg, A. L. (2004) Pathway for degradation of peptides generated by proteasomes: a key role for thimet oligopeptidase and other metallopeptidases. *J. Biol. Chem.* **279**, 46723–46732
  4. Gonzales, T., and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol. Rev.* **18**, 319–344
  5. Taylor, A. (1993) Aminopeptidases: toward a mechanism of action. *Trends Biochem. Sci.* **18**, 167–171
  6. Tasaki, T., Sriram, S. M., Park, K. S., and Kwon, Y. T. (2012) The N-end rule pathway. *Annu. Rev. Biochem.* **81**, 261–289
  7. van Endert, P. (2011) Post-proteasomal and proteasome-independent generation of MHC class I ligands. *Cell. Mol. Life Sci.* **68**, 1553–1567
  8. Matsui, M., Fowler, J. H., and Walling, L. L. (2006) Leucine aminopeptidases: diversity in structure and function. *Biol. Chem.* **387**, 1535–1544
  9. Guzman-Rojas, L., Rangel, R., Salameh, A., Edwards, J. K., Dondossola, E., Kim, Y. G., Saghatelian, A., Giordano, R. J., Kolonin, M. G., Staquicini, F. I., Koivunen, E., Sidman, R. L., Arap, W., and Pasqualini, R. (2012) Cooperative effects of aminopeptidase N (CD13) expressed by nonmalignant and cancer cells within the tumor microenvironment. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1637–1642
  10. Holz, R. C., Bzymek, K. P., and Swierczek, S. I. (2003) Co-catalytic metallopeptidases as pharmaceutical targets. *Curr. Opin. Chem. Biol.* **7**, 197–206
  11. Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., and Niedermann, G. (1999) A giant protease with potential to substitute for some functions of the proteasome. *Science* **283**, 978–981
  12. Remaut, H., Bompard-Gilles, C., Goffin, C., Frère, J. M., and Van Beeumen, J. (2001) Structure of the *Bacillus subtilis* D-aminopeptidase DppA reveals a novel self-compartmentalizing protease. *Nat. Struct. Biol.* **8**, 674–678
  13. Burley, S. K., David, P. R., Sweet, R. M., Taylor, A., and Lipscomb, W. N. (1992) Structure determination and refinement of bovine lens leucine aminopeptidase and its complex with bestatin. *J. Mol. Biol.* **224**, 113–140
  14. Joshua-Tor, L., Xu, H. E., Johnston, S. A., and Rees, D. C. (1995) Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. *Science* **269**, 945–950
  15. Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) Tricorn protease—the core of a modular proteolytic system. *Science* **274**, 1385–1389
  16. Du, X., Choi, I. G., Kim, R., Wang, W., Jancarik, J., Yokota, H., and Kim, S. H. (2000) Crystal structure of an intracellular protease from *Pyrococcus horikoshii* at 2-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14079–14084
  17. Delfosse, V., Girard, E., Birck, C., Delmarcelle, M., Delarue, M., Poch, O., Schultz, P., and Mayer, C. (2009) Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family. *PLoS One* **4**, e4712
  18. Franzetti, B., Schoehn, G., Hernandez, J. F., Jaquinod, M., Ruigrok, R. W., and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO J.* **21**, 2132–2138
  19. Schoehn, G., Vellieux, F. M., Asunción Durá, M., Receveur-Bréchet, V., Fabry, C. M., Ruigrok, R. W., Ebel, C., Roussel, A., and Franzetti, B. (2006) An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron. *J. Biol. Chem.* **281**, 36327–36337
  20. Talon, R., Kahn, R., Durá, M. A., Maury, O., Vellieux, F. M., Franzetti, B., and Girard, E. (2011) Using lanthanoid complexes to phase large macromolecular assemblies. *J. Synchrotron Radiat.* **18**, 74–78
  21. Rawlings, N. D., Barrett, A. J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **40**, D343–D350
  22. Russo, S., and Baumann, U. (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J. Biol. Chem.* **279**, 51275–51281
  23. Borissenko, L., and Groll, M. (2005) Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J. Mol. Biol.* **346**, 1207–1219
  24. Kim, D., San, B. H., Moh, S. H., Park, H., Kim, D. Y., Lee, S., and Kim, K. K. (2010) Structural basis for the substrate specificity of PepA from *Streptococcus pneumoniae*, a dodecameric tetrahedral protease. *Biochem. Biophys. Res. Commun.* **391**, 431–436
  25. Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K., and Kiser, P. D. (2012) Insights into substrate specificity and metal activation of mammalian tetrahedral aspartyl aminopeptidase. *J. Biol. Chem.* **287**, 13356–13370
  26. Chaikwad, A., Pilka, E. S., De Riso, A., von Delft, F., Kavanagh, K. L., Vénien-Bryan, C., Oppermann, U., and Yue, W. W. (2012) Structure of human aspartyl aminopeptidase complexed with substrate analogue: insight into catalytic mechanism, substrate specificity, and M18 peptidase family. *BMC Struct. Biol.* **12**, 14
  27. Durá, M. A., Receveur-Brechot, V., Andrieu, J. P., Ebel, C., Schoehn, G., Roussel, A., and Franzetti, B. (2005) Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry* **44**, 3477–3486
  28. Durá, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M., and Franzetti, B. (2009) The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol. Microbiol.* **72**, 26–40
  29. Rosenbaum, E., Ferruit, M., Durá, M. A., and Franzetti, B. (2011) Studies on the parameters controlling the stability of the TET peptidase superstructure from *Pyrococcus horikoshii* revealed a crucial role of pH and catalytic metals in the oligomerization process. *Biochim. Biophys. Acta* **1814**, 1289–1294
  30. Ando, S., Ishikawa, K., Ishida, H., Kawarabayashi, Y., Kikuchi, H., and Koguchi, Y. (1999) Thermostable aminopeptidase from *Pyrococcus horikoshii*. *FEBS Lett.* **447**, 25–28
  31. Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M., Moras, D., and Tarnus, C. (1994) Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure* **2**, 283–291
  32. Godfroy, A., Raven, N. D., and Sharp, R. J. (2000) Physiology and continuous culture of the hyperthermophilic deep-sea vent archaeon *Pyrococcus abyssi* ST549. *FEMS Microbiol. Lett.* **186**, 127–132
  33. Chamieh, H., Guetta, D., and Franzetti, B. (2008) The two PAN ATPases from *Halobacterium* display N-terminal heterogeneity and form labile complexes with the 20 S proteasome. *Biochem. J.* **411**, 387–397
  34. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619
  35. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* **80**, 2946–2953
  36. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864
  37. Wriggers, W. (2010) Using Situs for the integration of multiresolution structures. *Biophys. Rev.* **2**, 21–27
  38. Chacón, P., and Wriggers, W. (2002) Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.* **317**, 375–384
  39. Svergun, D., Barberato, C., and Koch, M. H. (1995) CRYSOLE—A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773
  40. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
  41. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996) SPIDER and WEB: processing and visualization of images in three-dimensional electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199
  42. Frotin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C., and Meinel, T. (2006) The proteomics of N-terminal methionine cleavage. *Mol. Cell. Proteomics* **5**, 2336–2349
  43. Ginzburg, M., Sachs, L., and Ginzburg, B. Z. (1970) Ion metabolism in a *Halobacterium*. I. Influence of age of culture on intracellular concentrations. *J. Gen. Physiol.* **55**, 187–207
  44. Glatter, O. (2002) *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*, p. 73, Elsevier, Amsterdam, The Netherlands
  45. Svedberg, T., Bauer, J. H., and Pedersen, K. O. (1940) *The ultracentrifuge*, Clarendon Press, Oxford
  46. Lupas, A., Flanagan, J. M., Tamura, T., and Baumeister, W. (1997) Self-compartmentalizing proteases. *Trends Biochem. Sci.* **22**, 399–404

## Assembling TET Peptidases

47. Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**, 463–471
48. Brandstetter, H., Kim, J. S., Groll, M., and Huber, R. (2001) Crystal structure of the tricorn protease reveals a protein disassembly line. *Nature* **414**, 466–470
49. O'Farrell, P. A., Gonzalez, F., Zheng, W., Johnston, S. A., and Joshua-Tor, L. (1999) Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease. *Structure* **7**, 619–627
50. Thompson, M. W., Miller, J., Maurizi, M. R., and Kempner, E. (1998) Importance of heptameric ring integrity for activity of *Escherichia coli* ClpP. *Eur. J. Biochem.* **258**, 923–928
51. Burley, S. K., David, P. R., Taylor, A., and Lipscomb, W. N. (1990) Molecular structure of leucine aminopeptidase at 2.7-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6878–6882
52. Marquardt, U., Zettl, F., Huber, R., Bode, W., and Sommerhoff, C. (2002) The crystal structure of human  $\alpha$ 1-tryptase reveals a blocked substrate-binding region. *J. Mol. Biol.* **321**, 491–502
53. Debela, M., Magdolen, V., Grimminger, V., Sommerhoff, C., Messerschmidt, A., Huber, R., Friedrich, R., Bode, W., and Goettig, P. (2006) Crystal structures of human tissue kallikrein 4: activity modulation by a specific zinc-binding site. *J. Mol. Biol.* **362**, 1094–1107
54. González, J. M., Masuchi, Y., Robb, F. T., Ammerman, J. W., Maeder, D. L., Yanagibayashi, M., Tamaoka, J., and Kato, C. (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2**, 123–130
55. Jaenicke, R., and Böhm, G. (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738–748
56. Clantin, B., Tricot, C., Lonhienne, T., Stalon, V., and Villeret, V. (2001) Probing the role of oligomerization in the high thermal stability of *Pyrococcus furiosus* ornithine carbamoyltransferase by site-specific mutants. *Eur. J. Biochem.* **268**, 3937–3942

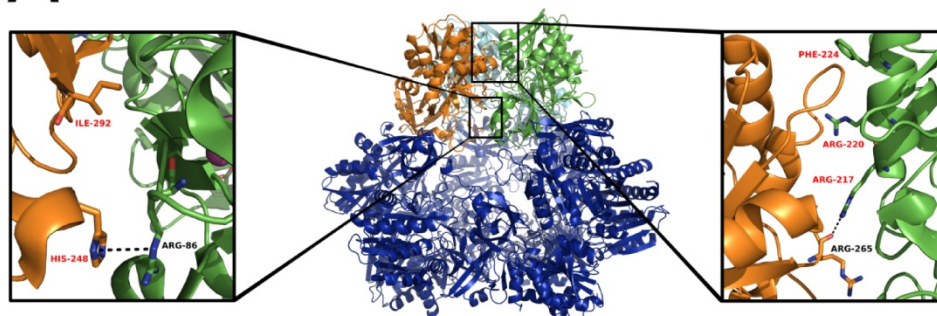
## Supplementary Material

**SUPPLEMENTARY FIGURE 1.** Positions of PhTET2 mutations and destabilized interfaces. Mutated residues are labelled in red, non-mutated interacting residues are labeled in black. A. Centre: Facet of the tetrahedral PhTET2 particle (PDB entry 1Y0R), three monomers forming one of the apices are coloured. Right: Close-up on the mutated apex interface; the mutations R217S, R220S and F224S are represented. Left: Close-up on the second mutated interface, the mutations H248S and I292A are represented. B. Right: View from one of the apex of the tetrahedron; three monomers forming the apex are coloured. Left: Close-up on the apex interface; residues of the blue monomer are labelled. C. Mutated residues and the corresponding eliminated interactions in the PhTET2 pentamutant. The destabilization of the oligomerisation interface is obtained by the elimination of hydrophobic and polar interactions.

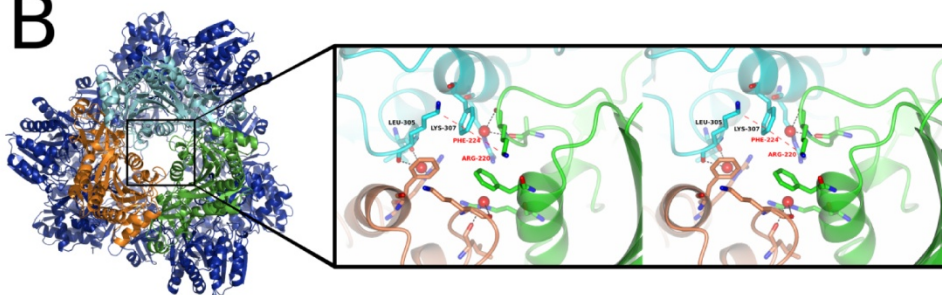


# **Supplementary Figure 1: Positions of PhTET2 mutations and destabilised interfaces.**

**A**



**B**



**C**

Mutation	Interaction type	Amino acid residue implicated	Description	Observations
R217S	Polar	R265	Main chain carbonyl oxygen	Water mediated
	Hydrophobic	D264	Main chain carbonyl oxygen	Mutation was done to S instead to A to completely avoid this interaction
		L305	Side chain	
R220S	Polar	L305	Main chain carbonyl oxygen	Water mediated
		K307	Main chain carbonyl oxygen	Water mediated
		F224S	F224 aromatic ring with side chain	Water mediated
H248S	Polar	E80	Carboxyl group	Water mediated
	Hydrophobic	H78	Side chain	Mutation was done to S instead to A to completely avoid these interactions
		R86		
I292A	Hydrophobic	T77	Side chain	The I292 CH does not interact. Therefore, mutation was done to A.
		I158		
		A88		

### 5.3 DISCUSSION ET PERSPECTIVES:

---

Le travail décrit dans cet article a mis en évidence l'existence de précurseurs dimériques de la peptidases TET dans les cellules. Il révèle que l'assemblage sert à activer la particule peptidasique en lui conférant une meilleure efficacité contre des longs polypeptides. Nos résultats indiquent également que des dimères thermostables, précurseurs du complexe TET, s'accumulent dans les cellules. Des résultats préliminaires du laboratoire chez *H. salinarum* ont montré que la formation de dodécamère est déclenchée *in vivo* dans certaines conditions métaboliques: début de phase stationnaire par exemple (résultats non publiés). Toutes ces observations suggèrent l'existence de mécanisme de régulation original *in vivo* portant sur un contrôle de l'oligomérisation des TET. Les dimères représenteraient une forme de stockage, faiblement active et mobilisable rapidement dans des conditions physiologiques qui demandent une activité de destruction des polypeptides. L'activité des TET serait régulée par un mécanisme stabilisateur du dimère qui pourrait être levé par un signal inconnu et des mécanismes restant à découvrir. Il faut toutefois prendre en considération le fait que les dimères de TET conservent une activité aminopeptidase contre les petits peptides. Cette activité pourrait être importante pour la cellule, soit pour la dégradation de tout petits substrats, soit pour modifier des extrémités N-ter de protéines accessibles. Ainsi une double fonction du complexe TET est-elle envisageable. Toutes ces hypothèses doivent maintenant être testées, notamment en étudiant plus finement l'état d'oligomérisation et d'association des peptidases TET dans les cellules et en tentant d'identifier des interactants susceptibles de stabiliser le dimère.

Il n'existe que peu d'information sur les mécanismes d'assemblage des structures quaternaires. Le travail réalisé sur TET présente donc un grand intérêt en biophysique fondamentale. En effet, la deuxième information importante que l'on peut retirer de notre étude est que l'assemblage des complexes TET s'opère d'une manière très ordonnée. Il ne s'agit pas d'une simple addition séquentielle de dimère. Le travail d'analyse de la cinétique d'apparition des différents intermédiaires, associé à un travail de modélisation et de description des propriétés biophysiques et structurales des intermédiaires, conduit à proposer que les interfaces contiennent une information qui "pilote" la séquence d'assemblage d'abord vers une forme hexamérique métastable,

forme intermédiaire qui permet ensuite la formation du dodécamère. Pour démontrer pleinement ce mécanisme, nous avons mis en œuvre une étude combinant la reconstruction *in vitro* de complexes TET partiellement deutérés, avec la diffusion neutronique aux petits angles (SANS). Ce travail est décrit dans le chapitre suivant.

## 6. ETUDE DE L'ASSEMBLAGE DE L'HETERO-OLIGOMERE DES AMINOPEPTIDASES TET2 ET TET3 DE *PYROCOCCUS HORIKOSHII* PAR DIFFUSION DE NEUTRONS AUX PETITS ANGLES.

### 6.1 PRESENTATION DU TRAVAIL

---

Nous avons vu dans le second chapitre que les aminopeptidases TET2 et TET3 de *Pyrococcus horikoshii* sont capables de former des hétéro-oligomères PhTET2:PhTET3 *in vivo*. De plus, l'analyse des propriétés enzymatiques d'un hétéro-oligomère produit par co-expression chez *E.coli* a montré que la forme iso-stœchiométrique de PhTET2:PhTET3 présente une activité enzymatique accrue.

Comme nous l'avons vu dans l'introduction, l'oligomérisation des peptidases joue un rôle décisif dans leur l'activité, pour le criblage des substrats comme pour la structuration des sites actifs, mais également dans leur stabilisation, enfin, elle est essentielle pour protéger le protéome de la cellule. Pourtant, le processus d'oligomérisation des grands assemblages en lui-même est souvent peu étudié. Dans le cas de complexes symétriques et homo-oligomériques, plusieurs difficultés peuvent expliquer cette lacune, la première est sans doute la difficulté à purifier et à stabiliser les intermédiaires d'oligomérisation pour leur caractérisation. La seconde est la difficulté à résoudre la position et l'arrangement des différents intermédiaires au sein du complexe qu'ils forment. Dans le chapitre précédent, j'ai pu venir à bout de la première difficulté afin de proposer un modèle pour l'oligomérisation de la peptidase TET.

Dans l'article présenté dans ce chapitre, nous avons utilisé la technique de diffusion de neutrons aux petits angles (SANS) afin d'étudier l'arrangement interne général des sous-unités du complexe PhTET2:PhTET3. L'idée de départ était de comparer les courbes expérimentales obtenues avec un hétéro-oligomère PhTET2:PhTET3 à des courbes théoriques calculées à partir d'une banque de modèles structuraux comprenant toutes les possibilités d'assemblages hétéro-oligomériques possibles à partir d'homo-dimère comme unités de base, cette dernière hypothèse est liée à la méthode de production de l'hétéro-oligomère *in vitro* présentée dans la suite. De cette manière nous pouvions

dévoiler la façon dont sont assemblées les différentes sous-unités au sein de l'hétéro-oligomère.

Pour réaliser cette étude, nous avons utilisé la technique de variation de contraste en SANS (voir section 2.1.3). Pour la mener à bien, la première étape était de mettre au point un protocole pour la production d'une particule hétéro-oligomérique PhTET2:PhTET3 dont seules les sous-unités PhTET2 seraient deutérées. La méthode classique de deutération d'une protéine consiste à faire croître les cellules d'*E. coli*, transformées avec le plasmide d'intérêt, dans un milieu enrichi en deutérium. La méthode de production d'hétéro-oligomère utilisée dans le chapitre 4 n'était donc pas adaptée, en effet, il n'aurait pas été possible de deutérer sélectivement une seule des deux protéines co-exprimées dans le même organisme. J'ai donc dû utiliser une méthode différente, moins directe, que je vais développer ici.

En collaboration avec Martine Moulin et Michael Härtlein du laboratoire de deutération (D-Lab) de l'Institut Laue-Langevin (ILL) nous avons pu produire et purifier la peptidase homo-oligomérique recombinante PhTET2 deutéré à 75 %. L'étape suivante a représenté une des grandes difficultés rencontrées dans ce projet.

Il a été montré que l'oligomérisation de TET est métal-dépendante (Rosenbaum *et col.*, 2011) (voir section 1.5). La stratégie que j'ai utilisée pour produire l'hétéro-oligomère PhTET2:PhTET3 bénéficie de cette particularité. Les protéines PhTET2 deutérées (notées dPhTET2) et PhTET3 hydrogénées (notées hPhTET3) sont purifiées, dialysées séparément contre un tampon de désoligomérisation contenant un chélateur d'ions métalliques (EDTA), puis regroupées avant que ne soit induite la réoligomérisation lors d'une dernière dialyse en utilisant un tampon enrichi en ions  $\text{Co}^{2+}$ . L'état d'oligomérisation des protéines est contrôlé entre les différentes étapes. Puis, des séries de chromatographies d'affinité et d'exclusion sont réalisées afin de purifier les différentes formes d'hétéro-oligomères produites.

Ces étapes représentent trois semaines d'expériences. Plusieurs essais ont été nécessaires pour optimiser le protocole et produire assez de protéines pour réaliser les expériences de SANS. Lors de ce processus, deux pics majoritaires de protéines ont été produits, et plusieurs autres minoritaires (au moins trois), seuls les deux pics

majoritaires contenaient assez de matériel pour réaliser les expériences de SANS. Les conductivités d'élutions de ces différentes formes étaient toutes intermédiaires à celles des formes homo-dodécamériques de PhTET2 et PhTET3. Des tests d'activités réalisés sur ces formes purifiées ont montré qu'elles présentaient les activités catalytiques de PhTET2 et de PhTET3. J'avais donc pu purifier des formes hétéro-oligomériques fonctionnelles de TET, reconstituées *in vitro*, et formées de sous-unités hydrogénées de hPhTET3 et de sous-unités deutérées de dPhTET2. Ces protéines ont donc été utilisées comme modèles dans l'étude qui suit.

Ensuite, trois séries d'expériences SANS ont été réalisées en juillet 2011, en juillet 2012 puis en mars 2013. Ce sont les expériences présentées dans ce chapitre. Les hétéro-oligomères hPhTET3:dPhTET2 reconstitués *in vitro* ont été préparés dans des tampons à quatre contrastes différents (sont indiquées entre parenthèses les sous-unités qui contribuent au signal enregistré à ces différents contrastes): 0% D<sub>2</sub>O (dPhTET2 et hPhTET3), 42% D<sub>2</sub>O (dPhTET2 seul), 70% D<sub>2</sub>O (dPhTET2 et hPhTET3, mais avec des signes opposés, voir section 2.1.3), et 100% D<sub>2</sub>O (hPhTET3 seul). L'ensemble des jeux de données a été enregistré sur la ligne D22 à l'ILL.

Avec l'aide d'Eric Girard et de Matteo Colombo, les structures cristallographiques de PhTET2 et PhTET3 ont été affinées par rapport aux données des expériences contrôles SANS réalisées sur des homo-oligomères. Ainsi, des boucles mobiles absentes des fichiers PDB (fichiers contenant les données cristallographiques), mais présentes en solution ont pu être modélisées. Afin de pouvoir analyser les données expérimentales obtenues, j'ai pu créer une banque de structures contenant toutes les possibilités d'assemblages hétéro-oligomériques à partir des structures cristallographiques ajustées de PhTET2 et PhTET3, avec l'hypothèse d'avoir un homo-dimère comme « building block ». Cette dernière hypothèse est posée car l'étude présentée dans le chapitre précédent montre qu'*in vivo*, il serait possible que se forment des complexes constitués d'hétéro-dimères. Cependant, *in vitro*, il est très difficile, si ce n'est impossible, de produire des hétéro-dimères sans dénaturer de façon importante la protéine. Le modèle pour l'assemblage de la particule TET reste valable, il doit cependant être nuancé avec la possibilité que se forment des hétéro-oligomères composés d'hétéro-dimères *in vivo*.



**6.2 SMALL ANGLE NEUTRON SCATTERING REVEALS THE ASSEMBLING  
MODE AND OLIGOMERIC ORGANIZATION OF TET, A LARGE  
DODECAMERIC AMINOPEPTIDASE**

---

# Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase

Alexandre Appolaire,<sup>a,b,c</sup> Eric Girard,<sup>a,b,c</sup> Matteo Colombo,<sup>a,b,c</sup> M. Asunción Durá,<sup>a,b,c</sup> Martine Moulin,<sup>d</sup> Michael Härtlein,<sup>d</sup> Bruno Franzetti<sup>a,b,c</sup> and Frank Gabel<sup>a,b,c,e\*</sup>

<sup>a</sup>Université Grenoble Alpes, IBS, 38044 Grenoble, France, <sup>b</sup>CNRS, IBS, 38044 Grenoble, France, <sup>c</sup>CEA, IBS, 38044 Grenoble, France, <sup>d</sup>Life Sciences Group, Institut Laue–Langevin, 38042 Grenoble CEDEX 9, France, and <sup>e</sup>Large Scale Structures Group, Institut Laue–Langevin, 38042 Grenoble CEDEX 9, France

Correspondence e-mail: frank.gabel@ibs.fr

Received 22 April 2014  
Accepted 13 August 2014

The specific self-association of proteins into oligomeric complexes is a common phenomenon in biological systems to optimize and regulate their function. However, *de novo* structure determination of these important complexes is often very challenging for atomic-resolution techniques. Furthermore, in the case of homo-oligomeric complexes, or complexes with very similar building blocks, the respective positions of subunits and their assembly pathways are difficult to determine using many structural biology techniques. Here, an elegant and powerful approach based on small-angle neutron scattering is applied, in combination with deuterium labelling and contrast variation, to elucidate the oligomeric organization of the quaternary structure and the assembly pathways of 468 kDa, hetero-oligomeric and symmetric *Pyrococcus horikoshii* TET2–TET3 aminopeptidase complexes. The results reveal that the topology of the *Ph*TET2 and *Ph*TET3 dimeric building blocks within the complexes is not casual but rather suggests that their quaternary arrangement optimizes the catalytic efficiency towards peptide substrates. This approach bears important potential for the determination of quaternary structures and assembly pathways of large oligomeric and symmetric complexes in biological systems.

## 1. Introduction

The specific self-association of proteins to form oligomeric machines is a common phenomenon in biological systems. Homo-oligomerization in a closed symmetry is particularly prevalent in enzymes and high-order oligomers comprising more than six subunits and represents a significant part of the proteomes (Matthews & Sunde, 2012). There are numerous reasons why enzymes self-assemble into large edifices: while representing a way to increase the stability and solubility of the system, the main advantage afforded by complex quaternary structures is to improve enzymatic function. Subunit interactions often allow the formation of channels to increase the specific affinity towards the substrate(s). Oligomerization also induces cooperativity between monomers to build up active sites or to share them at interfaces (Marianayagam *et al.*, 2004) and, in multi-subunit enzymatic complexes, the quaternary structure enhances functional cooperativity (Griffin & Gerrard, 2012). Monomers assembled into large, hollow edifices can create compartments for the confinement of biochemical activities. In the case of energy-dependent chaperonins or complexes of ATPases associated with various cellular activities (AAA-ATPases), this strategy is exploited to unfold, refold or disassemble macromolecular edifices in the interior of a nano-compartment, thus avoiding aggregation in the crowded cytosolic environment (Snider &

Houry, 2008). In the case of large intracellular peptidases such as proteasomes, the self-compartmentalization allows the confinement of the active sites inside the final particle and, in this way, uncontrolled proteolytic activity is avoided in the cytosol (Sauer & Baker, 2011). Finally, another advantage of oligomerization is to generate docking platforms for regulatory proteins or complexes. For all of these reasons, large enzymatic machines must adopt a well defined quaternary structure to carry out their function, and one way to regulate their activity within the cell is by controlling their oligomerization state.

Surprisingly, only a few studies have addressed the fundamental questions of the assembly pathways of large homooligomers. Progress has been made in recent years, indicating that dimeric precursors are often involved (Marianayagam *et al.*, 2004). However, the structural details and control mechanisms of the assembly processes are still largely unknown: are they orchestrated, stepwise processes with well defined oligomeric intermediates? Does assembly follow a limited number of pre-defined pathways or is it of a random nature? How are the intermediate oligomers positioned in the final complex? Concomitantly, the study of the assembly processes of homomeric, symmetric complexes poses two practical challenges: (i) the isolation, stabilization and biochemical/biophysical characterization of the building blocks and intermediate states and (ii) the determination of their respective positions and arrangement within the final edifice. Here, we apply a powerful and elegant approach based on the combination of small-angle neutron scattering (SANS), deuterium labelling and contrast variation to elucidate the oligomeric organization of the quaternary structures and the assembly modes of symmetric, heterododecameric, 468 kDa TET aminopeptidase complexes.

TET complexes are bi-metallic aminopeptidases that act as peptide-destruction machines and are widespread in the three domains of life. They belong to the M42 and M18 families in clan MH according to the MEROPS classification system (Rawlings *et al.*, 2014). The typical TET tetrahedral quaternary structure was initially described in archaea (Franzetti *et al.*, 2002; Russo & Baumann, 2004; Borissenko & Groll, 2005), but has also been found in bacteria (Kim *et al.*, 2010) and in eukarya (Chen *et al.*, 2012; Chaikuad *et al.*, 2012). Unlike most homomers, which adopt cyclic or dihedral symmetries, TET peptidases display unusually sophisticated quaternary structures, with 12 monomers being arranged in a hollow, tetrahedral edifice (Franzetti *et al.*, 2002). The crystallographic structures and the enzymatic properties of three different complexes (*PhTET1*, *PhTET2* and *PhTET3*) from the hyperthermophilic archaeon *Pyrococcus horikoshii* have been determined (Durá *et al.*, 2005, 2009; Schoehn *et al.*, 2006). They form self-compartmentalized assemblies with a common internal organization made up of a network of four access channels extended by four vast catalytic chambers, each containing three active sites close to the tetrahedron apices (Durá *et al.*, 2005). Each TET subunit is formed by a proteolytic and a PDZ-like domain that mediate both the dimeric interface and the dimer–dimer interface that build the

whole tetrahedral particle. We recently demonstrated that the activity of TET aminopeptidase towards long polypeptides is coupled with its assembly process (Appolaire *et al.*, 2013; Rosenbaum *et al.*, 2011): the co-occurrence *in vivo* of stable TET dimeric precursors and assembled dodecamers suggests that, at least in archaea, the TET oligopeptidase activity is regulated by control of its oligomerization state. *PhTET* homododecamers are very stable, even at high temperatures (90°C), a fact attributed to the properties of the dimer interfaces (Appolaire *et al.*, 2013). *In vivo* studies, combined with small-angle X-ray scattering (SAXS), showed that dimers are the elementary building blocks and, based on electron microscopy (EM) and analytical ultracentrifugation (AUC) data, a working model with hexameric intermediates assembling into the final dodecameric edifices has been suggested (Appolaire *et al.*, 2013).

The three M42 protein homomers comprised in the genome of *P. horikoshii* display complementary substrate specificities: *PhTET1* is a glutamyl-aminopeptidase that is active towards acidic residues, *PhTET2* is a leucyl-aminopeptidase that is active towards neutral residues, and *PhTET3* is a lysyl-aminopeptidase that is mainly active towards basic residues. The three forms therefore represent an integrated cellular peptide-destruction system (Durá *et al.*, 2009). *PhTET2* and *PhTET3* display a high sequence identity, are robust thermozymes and their quaternary structures resist harsh physicochemical treatments (Rosenbaum *et al.*, 2012). Their assembly process is metal-dependent and, *in vitro*, addition of a metal chelator agent combined with high-pH conditions is necessary to disassemble the complex into its stable dimeric precursors. These can spontaneously reform the tetrahedral edifice upon dialysis against a physiological assembly buffer containing cobalt (Rosenbaum *et al.*, 2011).

Here, we show that when mixed together in the assembly buffer *PhTET2* and *PhTET3* dimers can self-organize into 468 kDa heterododecameric complexes that possess the same quaternary structure as homododecameric TETs. Using a combination of SANS and contrast variation (variable H<sub>2</sub>O/D<sub>2</sub>O ratio in the solvent; Jacrot, 1976) on heterododecameric complexes of deuterated *PhTET2* (*dPhTET2*) and hydrogenated *PhTET3* (*hPhTET3*) allowed us to elucidate their respective quaternary architectures within the assembled hetero-complex. Our results demonstrate that TET complexes are built following a very limited number of specific and well defined pathways. Intriguingly, the resulting geometric arrangement of the two different building blocks within the heterododecameric complex is not casual but rather suggests a mechanism to optimize its catalytic properties towards peptide substrates. Our approach represents an elegant and attractive method to address structural questions in other challenging oligomeric systems composed of symmetric (or very similar) building blocks such as the proteasome. Finally, a better structural insight into the assembly mechanisms of large complexes is essential for preparing studies of the underlying regulatory mechanisms, which have not been intensively explored to date but might represent attractive targets for biomedical approaches and drug design.

## 2. Material and methods

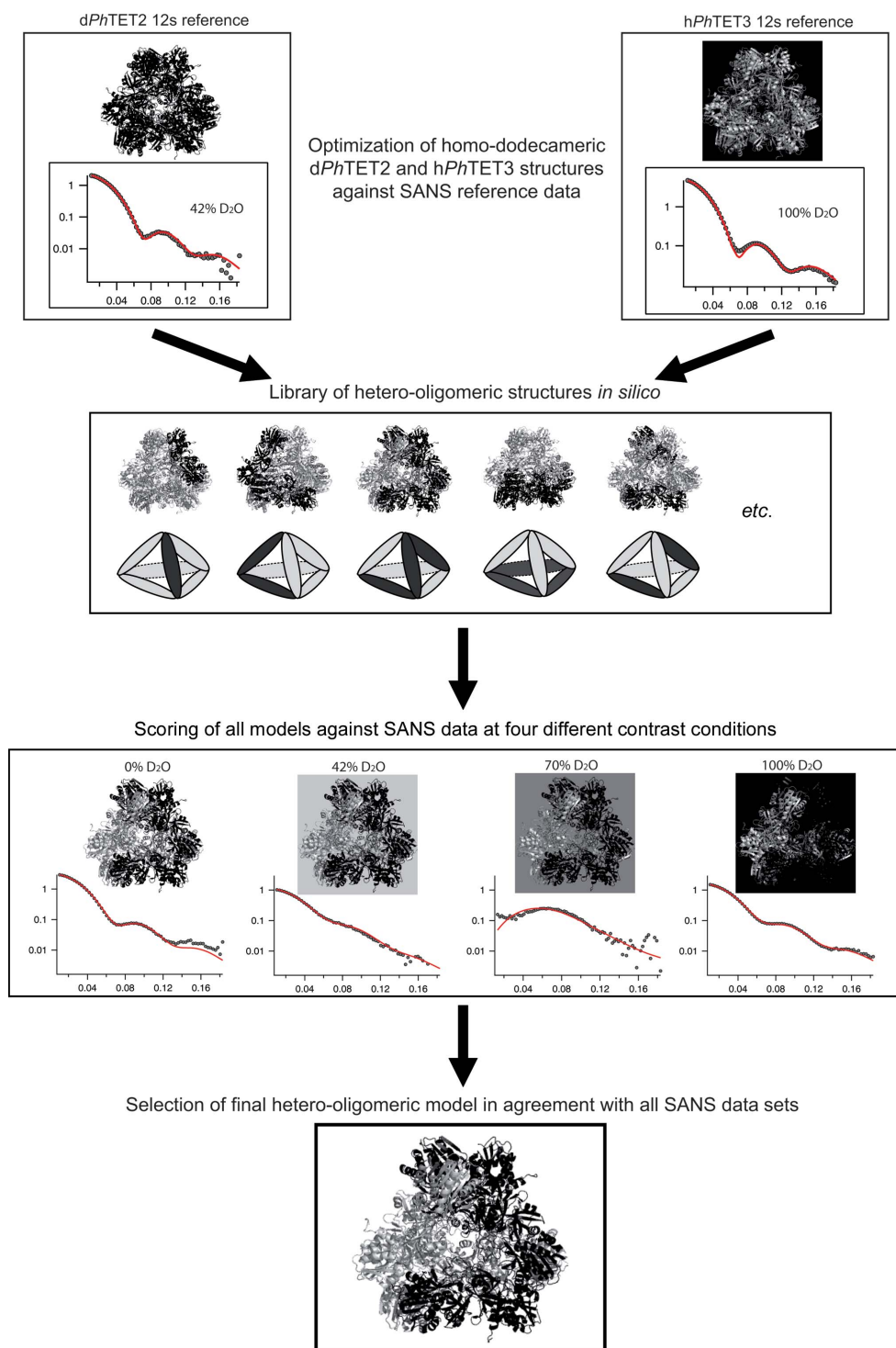
### 2.1. Expression and purification of hydrogenated *PhTET3* and deuterated *PhTET2*

Total protein extracts of *Escherichia coli* cells expressing the various recombinant hydrogenated *PhTET3* proteins

(*hPhTET3*) were purified as described by Durá *et al.* (2009). For SANS experiments, random-fractional deuteration of the wild type and a pentamutant (see below) *PhTET2* protein (*dPhTET2*) was carried out in the ILL Deuteration Laboratory, Grenoble, France. Cells were grown at 30°C in minimal medium as described by Artero *et al.* (2005) using 85%(v/v)

D<sub>2</sub>O and unlabelled glycerol as a carbon source. At an OD<sub>600</sub> of about 10 for the wild type and of about 13 for the pentamutant construct, the high cell-density cultures were induced with 1 mM IPTG overnight. The final deuteration level was approximately 75% to yield a SANS contrast match point of 100% D<sub>2</sub>O. The resulting cell pellets were stored at −80°C until further use.

For purification, the pellets were thawed at room temperature and resuspended in 50 ml 50 mM Tris–HCl, 150 mM NaCl, 0.1% Triton X-100 pH 8.0. Next, 12.5 mg lysozyme (Euromedex), 2.5 mg DNase I grade II (Roche), 10 mg RNase (Roche), 50 mg Pefabloc SC (Roche) and 0.5 ml 2 M MgSO<sub>4</sub> were added to the cell suspensions. Disruption of the cells was achieved by sonication in a Branson Sonifier 150 at 4°C. Five 30 s bursts at intensity 10 with intermediary pauses of 30 s were employed. The crude extract was then heated at 85°C for 15 min to eliminate most of the mesophilic proteins from the *E. coli* host and the lysates were clarified by centrifugation at 17 400g for 1 h. Supernatant concentrations were adjusted to 100 mM NaCl (*dPhTET2*) or 250 mM NaCl (*hPhTET3*), 20 mM Tris pH 7.5. The resulting extracts were loaded onto a 6 ml Resource Q column (GE Healthcare) previously equilibrated in 100 mM NaCl, 20 mM Tris–HCl pH 7.5. After washing with three column volumes (CV) of the equilibration buffer, bound proteins were eluted at 4 ml min<sup>−1</sup> with a 20 CV linear salt gradient (0.1–0.35 M NaCl for *dPhTET2* or 0.25–0.45 M NaCl for *hPhTET3* in 20 mM



**Figure 1**

Workflow of the SANS approach adopted to determine the quaternary arrangement of deuterated *PhTET2* (black) and hydrogenated *PhTET3* (grey) building blocks within the heterododecameric particles in solution.

	Sample	6s6s 'Z'	6s6s 'tripod-triangle'	4s8s 'opposite'	8s4s 'opposite'	4s8s 'V'	8s4s 'V'	2s10s	10s2s
Peak 1	0% D <sub>2</sub> O	2.93	5.38	3.31	1.84	5.02	3.63	3.46	3.58
	42% D <sub>2</sub> O	0.99	3.36	5.29	5.43	3.35	4.45	10.52	7.92
	70% D <sub>2</sub> O	1.78	6.15	1.20	1.13	4.81	5.54	7.19	9.26
	100% D <sub>2</sub> O	4.08	15.22	1.60	19.02	10.04	20.12	12.18	33.00
Peak 2	0% D <sub>2</sub> O	5.40	7.18	6.53	3.17	7.92	5.32	6.37	4.24
	49% D <sub>2</sub> O	2.17	6.27	14.47	0.95	12.30	4.12	26.41	6.96
	70% D <sub>2</sub> O	0.73	2.66	1.45	1.44	1.86	2.02	2.92	4.64
	100% D <sub>2</sub> O	6.31	15.67	2.43	1.64	12.35	13.43	16.60	16.32

**Figure 2**

Overview of the fits of several homodimeric models against the four SANS contrast scattering curves. Deuterated *PhTET2* dimers are depicted by dark grey ellipsoids and hydrogenated *PhTET3* dimers are depicted as light grey ellipsoids. Green indicates excellent fits against SANS data at the respective H<sub>2</sub>O/D<sub>2</sub>O contrast, yellow moderate agreement and red strong disagreement. The numbers indicate the  $\chi^2$  values obtained with *CRYSON*. Peak 1 can only be fitted by one model in a satisfactory way (*dPhTET2* 6s:*hPhTET3* 6s = 'Z'), while peak 2 can only be fitted by a single, distinct model (*dPhTET2* 8s:*hPhTET3* 4s = 'opposite'). All other dodecameric models based on homodimeric building blocks can readily be ruled out.

Tris–HCl pH 7.5). For further purification, the proteins were loaded onto a Superose 6 size-exclusion column (GE Healthcare) equilibrated and run in 20 mM Tris–HCl, 150 mM NaCl pH 7.5.

## 2.2. De-oligomerization of native TET homododecamers and formation of heterododecamers

After the size-exclusion step of the purification, dodecameric *dPhTET2* and *hPhTET3* samples were dialyzed against a de-oligomerization buffer (50 mM CAPS, 20 mM NaCl, 20 mM EDTA pH 10). Aliquots of each sample were analyzed by size exclusion on a Superose 6 column (GE Healthcare) and on native polyacrylamide gels to control the oligomeric state of the samples after dialysis. According to the size-exclusion chromatogram (Supplementary Fig. S1<sup>1</sup>), both contained mostly TET dimers; the *dPhTET2* sample contained less than 4% monomers and the *hPhTET3* sample contained approximately 30% monomers. Therefore, less than 4% of the dimers present in solution can be *dPhTET2*–*hPhTET3* heterodimers and such building blocks were therefore discarded in the modelling process, which was exclusively based on homodimeric building blocks (Figs. 1 and 2).

The de-oligomerized *dPhTET2* and *hPhTET3* samples were mixed together and dialyzed against re-oligomerization buffer (20 mM Tris–HCl, 150 mM NaCl, 2 mM CoCl<sub>2</sub> pH 7.5). The re-oligomerized sample was analyzed on a Resource Q ion-exchange column (GE Healthcare) equilibrated in 150 mM NaCl, 20 mM Tris–HCl pH 7.5. After washing with 3 CV of the equilibration buffer, bound proteins were eluted at 4 ml min<sup>−1</sup> with a 20 CV linear salt gradient (0.15–0.45 M

NaCl in 20 mM Tris–HCl pH 7.5). Four peaks were obtained, two of which were predominant and contained a sufficient amount of protein for further characterization by SANS. According to their order of elution during the ion-exchange chromatography step, they were named peak 1 and peak 2. To increase the monodispersity of the samples, they were loaded onto a 1 ml Mono Q column (GE Healthcare) equilibrated in 150 mM NaCl, 20 mM Tris–HCl pH 7.5. After washing with 3 CV of the equilibration buffer, bound proteins were eluted at 1.5 ml min<sup>−1</sup> with a 30 CV linear salt gradient (0.15–0.45 M NaCl in 20 mM Tris–HCl pH 7.5). They were then analyzed by size-exclusion chromatography (Superose 6) to control their final oligomerization state and both were exclusively dodecameric (Supplementary Fig. S10).

## 2.3. SANS sample preparation, experimental details and data reduction

The samples corresponding to each peak from the final size-exclusion chromatography were split into two: one half was conserved in 100% H<sub>2</sub>O and the other half was equilibrated in 100% D<sub>2</sub>O. Buffer exchange (20 mM Tris–HCl, 150 mM NaCl pH 7.5 in 100% D<sub>2</sub>O or 100% H<sub>2</sub>O) was performed by dialysis overnight at 4°C. The D<sub>2</sub>O ratio of the different samples was then obtained by mixing the two solutions. 200 µl of protein solution was prepared at a concentration of 4.5 mg ml<sup>−1</sup> at the following D<sub>2</sub>O ratios: 0, 42, 70 and 100% D<sub>2</sub>O. The D<sub>2</sub>O ratios were validated *via* their neutron transmission values. Apart from the 42% sample of peak 2, which was actually at 49% D<sub>2</sub>O, all samples were at the nominal ratio. Homododecameric *hPhTET2* (42% D<sub>2</sub>O), *dPhTET2* (42 and 100% D<sub>2</sub>O) and *hPhTET3* (100% D<sub>2</sub>O) reference samples were prepared following the same protocol.

<sup>1</sup> Supporting information has been deposited in the IUCr electronic archive (Reference: WA5075).

**Table 1**

Model-free SANS parameters of the heterododecameric and homododecameric complexes at 20°C.

The forward scattered intensities  $I(0)$  and radii of gyration  $R_g$  extracted from the Guinier ranges ( $Q_{\min}R_g - Q_{\max}R_g$ ; fits are shown as an inset in Fig. 4a) are shown in comparison to the values extracted by indirect Fourier transform using *GNOM* (Svergun, 1992). The values extracted from both approaches are in very good agreement. The Guinier parameters were not determined for the 70% D<sub>2</sub>O data sets owing to high noise levels. The two *GNOM*  $R_g$  values represent the direct/indirect-space values, respectively.

Sample	$I(0)$ (Guinier)	$R_g$ (Guinier) (Å)	$Q_{\min}R_g - Q_{\max}R_g$	$R_g$ ( <i>GNOM</i> ) (Å)	$D_{\max}$ ( <i>GNOM</i> ) (Å)
Peak 1 (0% D <sub>2</sub> O)	3.28 ± 0.01	48.6 ± 0.2	0.77–1.40	47.8/48.0	130 ± 5
Peak 1 (42% D <sub>2</sub> O)	0.56 ± 0.01	45.4 ± 0.7	0.72–1.31	45.8/45.8	130 ± 5
Peak 1 (70% D <sub>2</sub> O)	N.D.	N.D.	N.D.	34.4/38.1	130 ± 5
Peak 1 (100% D <sub>2</sub> O)	1.66 ± 0.01	47.6 ± 0.2	0.75–1.37	47.0/47.2	125 ± 5
Peak 2 (0% D <sub>2</sub> O)	3.84 ± 0.02	48.6 ± 0.2	0.77–1.40	48.1/48.3	130 ± 5
Peak 2 (49% D <sub>2</sub> O)	0.62 ± 0.01	47.7 ± 0.6	0.75–1.38	47.5/47.6	130 ± 5
Peak 2 (70% D <sub>2</sub> O)	N.D.	N.D.	N.D.	24.3/25.6	130 ± 5
Peak 2 (100% D <sub>2</sub> O)	0.90 ± 0.01	48.6 ± 0.3	0.65–1.40	47.3/47.5	125 ± 5
dPhTET2 12s (42% D <sub>2</sub> O)	2.32 ± 0.01	49.6 ± 0.3	0.78–1.43	48.4/48.6	130 ± 5
hPhTET3 12s (100% D <sub>2</sub> O)	5.28 ± 0.02	50.9 ± 0.2	0.68–1.33	49.1/49.3	135 ± 5

Protein concentrations were measured using the Bio-Rad protein-assay reagent (Bio-Rad) with bovine serum albumin as a standard. Correction factors were applied to pure *PhTET2* and *PhTET3* samples. These factors were calculated after determining the protein concentration of pure *PhTET2* and *PhTET3* samples by quantitative amino-acid analysis as described in Durá *et al.* (2009). For the concentration of pure hetero-oligomeric samples, a weighted correction factor was calculated taking into account the correction factors obtained for *PhTET2* and *PhTET3* and the number of monomers of these proteins in the heteromeric complexes. Therefore, all of the values reported in this paper refer to real protein concentrations. We estimate the accuracy of the concentration measurements to be about 30%.

SANS experiments were performed on the large dynamic range small-angle diffractometer D22 at the Institut Laue–Langevin (ILL), Grenoble, France. The incident wavelength was  $\lambda = 6$  Å ( $\Delta\lambda/\lambda = 10\%$ ) at a single detector/collimator configuration (4 m/4 m) with a centred beam. All samples and buffers were prepared the night before the experiment as described above. The final sample volumes were adjusted to 160 µl and placed in Hellma 1 mm QS quartz cells. Boron, empty cell and a pure H<sub>2</sub>O sample were measured as references for the subsequent data treatment. In a first experiment, we verified the match points of dPhTET2 in 100% D<sub>2</sub>O and hPhTET2 in 42% D<sub>2</sub>O (Supplementary Fig. S9). The temperature of the remaining samples, including buffers and references, was increased in steps of 20°C from 20 to 80°C. Typical exposure times were 20 min for the reference samples, mixtures and buffers at each temperature step. In addition, the empty cell, H<sub>2</sub>O and boron were measured for 10 min at each temperature. As a result, all samples were kept at each temperature for 4 h. Transmissions were measured for 3 min at the lowest temperature (20°C). The two-dimensional data sets of all samples and buffers were reduced to one-dimensional curves using the standard ILL software (Gosh *et al.*, 2006). Buffer intensities were subtracted from the respective sample intensities using *PRIMUS* from the *ATSAS* program suite (Konarev *et al.*, 2003).

## 2.4. SANS data analysis and modelling

The model-free parameters  $I(0)$  (intensity scattered in the forward direction) and  $R_g$  (radius of gyration) were extracted from the SANS curves using the Guinier approximation (Guinier, 1939; Table 1). Pair-distance distribution functions  $p(r)$  were extracted using *GNOM* (Svergun, 1992) by imposing  $p(r=0) = p(r=D_{\max}) = 0$ . Calibration of the molecular mass of the particles in solution was performed against water (Jacrot & Zaccai, 1981). The solvent-excluded volumes and scattering lengths of dPhTET2 and hPhTET3 monomers were calculated based on their amino-acid sequences (Jacrot, 1976), measured  $I(0)$ , protein concentrations, transmissions and quartz-cell path lengths (Supplementary Table S1).

*Ab initio* shapes of the *PhTET2* and *PhTET3* moieties within the heterododecamers (Supplementary Fig. S7) were calculated with *MONSA* (Svergun, 1999) using two phases (dPhTET2 and hPhTET3) according to the following parameters: monomer volumes  $V_{\text{PhTET2}} = V_{\text{PhTET3}} = 50\,000$  Å<sup>3</sup> (6s:6s stoichiometry, where 's' designates a single subunit, *i.e.* a monomer),  $V_{\text{PhTET2}} = 66\,700$  Å<sup>3</sup>,  $V_{\text{PhTET3}} = 33\,300$  Å<sup>3</sup> (8s:4s stoichiometry); the contrasts of dPhTET2 were 5.75, 3.34 (2.91) and  $-0.02 \times 10^{10}$  cm<sup>-2</sup> in 0, 42 (49) and 100% D<sub>2</sub>O and the contrasts of hPhTET3 were 2.30,  $-0.12$  ( $-0.53$ ) and  $-3.48 \times 10^{10}$  cm<sup>-2</sup> in 0, 42 (49) and 100% D<sub>2</sub>O. Connectivity of the two phases was not imposed.

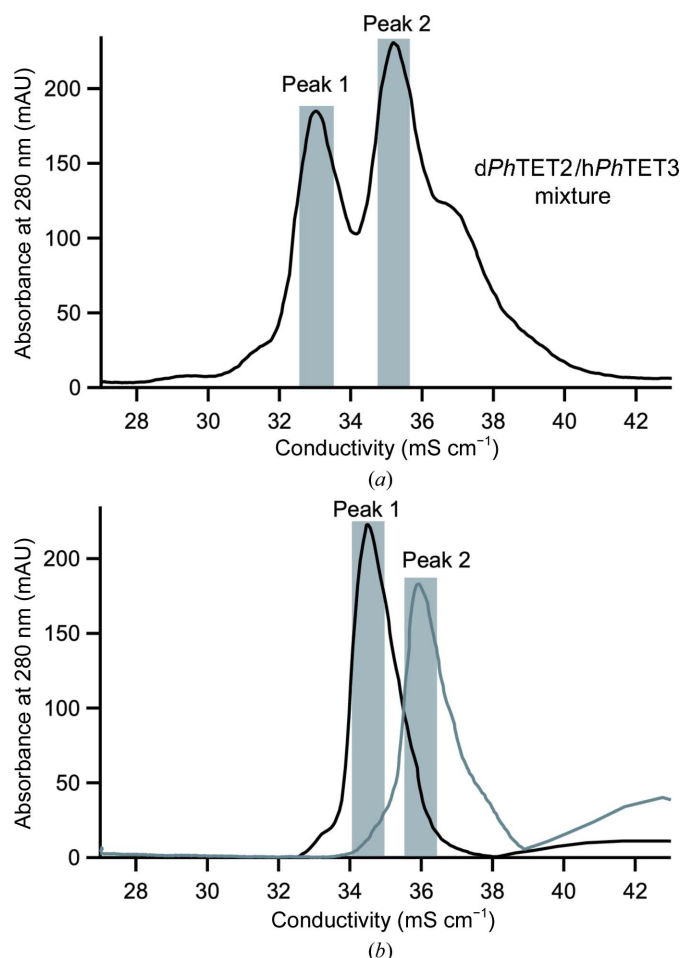
For the pseudo-atomic models, *PhTET2* and *PhTET3* dodecameric models were generated from the crystal structures (PDB entries 1y0r and 2wzn, respectively; Borissenko & Groll, 2005; Durá *et al.*, 2009). Indeed, both crystal structures present a monomer within the asymmetric unit, so generating the dodecameric models relied on only the same symmetry operators. Both dodecameric models were adjusted to yield best fits (Fig. 1, top; Supplementary Fig. S12) against the respective SANS curves of the homododecameric reference samples dPhTET2 12s (42% D<sub>2</sub>O) and hPhTET3 12s (100% D<sub>2</sub>O) using *CRYSON* (Svergun *et al.*, 1995, 1998). In particular, the missing parts of the respective crystal structures were added manually with *Coot* (Emsley & Cowtan, 2004). In the case of the *PhTET2* model, addition of the missing part



of the N-terminus (residues 1–5) as well as the internal loop (residues 120–132) led to a good fit of the d*PhTET2* 12 s SANS reference curve in 42% D<sub>2</sub>O (Fig. 1, top; Supplementary Fig. S12). As a basic unit, dimeric building blocks were chosen (in agreement with previous SAXS data; Appolaire *et al.*, 2013) and generated with the symmetry operator prior to generating the whole dodecameric model. In the case of *PhTET3*, addition of the N-terminal part (residues 1–7) as well as the internal loop (residues 128–136) led to a non-satisfactory fit of the h*PhTET3* 12 s SANS reference curve in 100% D<sub>2</sub>O. Consequently, the *PhTET3* dimeric building block was generated using a manual rigid body to slightly change the relative position of one monomer with respect to the second monomer within the dimeric building block. In this way and by adding the N-terminal part, a dodecameric model was generated with a reasonably good fit to the SANS reference curve (Fig. 1, top; Supplementary Fig. S12). An overview comparing both of the dodecameric TET particles before and after

the modifications as well as with each other is provided in Supplementary Fig. S11.

In order to facilitate the generation of heterododecameric hybrids, prior to dodecamer generation *PhTET2* and *PhTET3* dimeric models were superimposed in *Coot*. A library of heterododecameric hybrids was then created by editing the PDB files. The library exhaustively covered all dodecameric models based on homodimeric *PhTET2* and *PhTET3* building blocks (Fig. 2). All structures from the library were scored in a least  $\chi^2$  fit against the SANS data in 0, 42 (49), 70 and 100% D<sub>2</sub>O for peak 1 and peak 2 using *CRYSON*. Fits were classified as excellent, moderate and bad for each model and contrast (Fig. 2) and were colour-coded green, yellow and red, respectively. A complete version of Fig. 2, including the individual fit curves, is provided as Supplementary Fig. S13. It illustrates that both  $\chi^2$  values and visual criteria (good overall superposition of back-calculated and experimental curves) should be considered when discriminating between different models.



**Figure 3**  
Purification of heterododecamers. The peaks are represented using the absorbance ( $A_{280\text{ nm}}$ ) as a function of the conductivity ( $\text{mS cm}^{-1}$ ). (a) Ion-exchange chromatography (Resource Q) of the total sample after the re-oligomerization process. The shaded areas correspond to the fractions grouped together for successive Mono Q purification. (b) Ion-exchange chromatography (Mono Q) of the two major peaks after the first Resource Q step. The shaded areas correspond to the fractions grouped together for size-exclusion chromatography on a Superose 6 column (Supplementary Fig. S10).

### 3. Results

#### 3.1. Purification of *PhTET2*–*PhTET3* hetero-oligomers

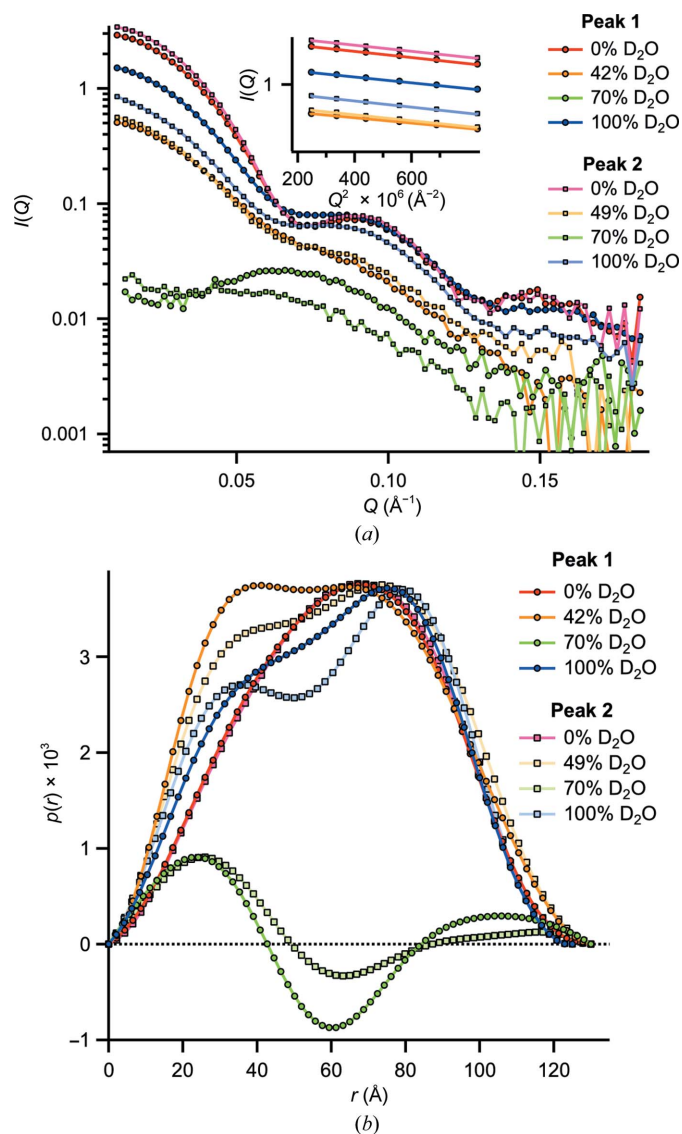
In the present study, we induced de-oligomerization of dodecameric deuterium-labelled *PhTET2* ('d*PhTET2*') and dodecameric hydrogenated *PhTET3* ('h*PhTET3*') into dimers. Subsequently, hetero-oligomers were generated by mixing the two dimer populations and inducing their re-oligomerization. Since *PhTET2* and *PhTET3* do not have the same external surface charges, their elution volumes are very different when analyzed appropriately by ion-exchange chromatography. Consequently, the different hetero-oligomeric assemblies can be separated by using an adequate ion-exchange chromatographic column run with a suitable salt gradient. After a first chromatographic step on a Resource Q column, two major (and several minor) peaks were observed (Fig. 3a). In order to better separate the different hetero-oligomers, a second strong ion-exchange chromatography was performed on a Mono Q column (Fig. 3b). (The fractions recovered from the Resource Q for the Mono Q are shaded in grey in Fig. 3a.) The protein concentrations and monodispersity of the two major peaks ('peak 1' and 'peak 2') were high enough for subsequent SANS experiments.

#### 3.2. SANS data reveal two extremely stable, heterododecameric d*PhTET2*–h*PhTET3* architectures of different stoichiometry and quaternary structure

The samples from both peaks yielded very specific and distinct SANS curves (Fig. 4a) and model-free parameters  $I(0)$  and  $R_g$  (Table 1) at the four contrast conditions. At 0% D<sub>2</sub>O both resemble the homododecameric reference curves (Supplementary Fig. S6) most since both d*PhTET2* and h*PhTET3* moieties have positive contrast with respect to the solvent (Supplementary Fig. S4). The 70% D<sub>2</sub>O data have the lowest relative intensities since the hydrogenated and deuterated building blocks are of opposite contrast and result in a

decrease in the overall scattered intensities. At 42 (49) and 100% D<sub>2</sub>O the scattered signals are almost exclusively owing to the deuterated and hydrogenated building blocks of the particles, respectively. Since small but significant variations of the curves at 0, 42 (49), 70 and 100% D<sub>2</sub>O are observed for both peaks, it can be concluded that the geometric arrangement of dPhTET2 and hPhTET3 building blocks are different within the particles constituting peak 1 and peak 2.

An intriguing feature is the evolution of the  $I(0)$  intensities as a function of contrast: while particle 2 (peak 2) scatters



**Figure 4**

SANS curves and  $p(r)$  functions of peak 1 and peak 2. (a) SANS curves at four different contrast conditions at 20°C. All data sets are drawn without applying scaling. The total protein concentrations of all samples were identical (4.5 mg ml<sup>-1</sup>). Guinier fits to the 0, 42 (49) and 100% D<sub>2</sub>O data are shown as an inset (the 70% D<sub>2</sub>O data were not fitted using the Guinier approach owing to their elevated noise level). (b) Pair-distance distribution functions  $p(r)$  of the SANS data in arbitrary units generated with *GNOM* (Svergun, 1992). The two 0% D<sub>2</sub>O data sets are very similar (*cf.* Supplementary Fig. S6). The 0, 42, 49 and 100% D<sub>2</sub>O data are normalized to the second peak and the 70% D<sub>2</sub>O data to the first peak (using a different scaling factor for clarity).

more strongly in the forward direction at 0, 42 and 70% D<sub>2</sub>O than particle 1 (peak 1), the situation is the inverse at 100% D<sub>2</sub>O. Since  $I(0)$ , at a given total protein concentration, is proportional to the contrast integrated over the whole particle volume (Svergun & Koch, 2002), these findings indicate that particle 2 contains more deuterated building blocks than particle 1. Interestingly, these findings imply that at least two stoichiometrically different architectures are possible for heterododecameric TET particles *in vitro*. A calibration of the molecular masses against water (Jacrot & Zaccari, 1981) of the 0% D<sub>2</sub>O data yielded a dPhTET2:hPhTET3 stoichiometry of 6:6 for particle 1 and of 8:4 for particle 2 (Supplementary Table S1). Importantly, both hetero-oligomeric TET particles display extreme structural stability at high temperatures, with the SANS curves being virtually identical at 20 and 80°C (Supplementary Fig. S2). The same stability was found for mixtures of preformed homododecameric dPhTET2 and hPhTET2 particles for both the wild type and a pentamutant variant designed to weaken the oligomerization interface (Supplementary Fig. S3). Importantly, these data rule out a dynamic equilibrium between dodecameric assemblies and putative smaller oligomers in solution under our experimental conditions. The quality of the matching of hydrogenated and deuterated PhTET2 was checked by measuring dodecameric hPhTET2 and dPhTET2 particles in 42 and 100% D<sub>2</sub>O, respectively. The results (Supplementary Fig. S9) indicated that the matching conditions were excellent, *i.e.* contributions from hydrogenated moieties can be neglected at 42% D<sub>2</sub>O and those of deuterated moieties at 100% D<sub>2</sub>O.

The pair-distance distribution functions  $p(r)$  extracted with *GNOM* (Svergun, 1992) from the SANS curves at the four contrast conditions revealed a wealth of real-space information on the relative positions of deuterated and hydrogenated building blocks within the two particles (Fig. 4b). The back-fitted (regularized) scattering curves are shown in Supplementary Fig. S5 and show excellent agreement with the experimental data. An overview of the numerical values of  $R_g$  and  $D_{\max}$  is provided in Table 1.

At 0% D<sub>2</sub>O both dPhTET2 and hPhTET3 building blocks have positive contrast (of different amplitude) with respect to the solvent (Supplementary Fig. S4). The  $p(r)$  functions from peak 1 and peak 2 are therefore almost identical, indicating that both particles possess a very similar overall shape (envelope) comparable to the dPhTET2 and hPhTET3 homododecameric references (Supplementary Fig. S6). The maxima of the leaning bell curves are shifted from the centre to larger distances and then drop rapidly to zero (a hallmark of hollow, globular particles; Koch *et al.*, 2003) and are in good agreement with the overall topology of dodecameric TET particles. The  $D_{\max}$  of 130 Å is also in excellent agreement with the expected PhTET dimensions from electron microscopy (Schoehn *et al.*, 2006).

At 42 (49)% D<sub>2</sub>O the  $p(r)$  functions are almost exclusively owing to deuterated moieties and therefore represent the internal arrangement of dPhTET2 building blocks within the dodecameric edifices. In both cases a broadened central peak is observed. A tendency to split up into a bimodal pattern

is observed for both particles: symmetric for particle 1 and asymmetric for particle 2. The observed bimodal  $p(r)$  patterns can be interpreted in terms of several compact, deuterated moieties positioned at a finite distance from each other, being either loosely connected or separated by small gaps. The 100% D<sub>2</sub>O data represent the complementary contrast condition, now with the hydrogenated building blocks (hPhTET3) almost exclusively accounting for the  $p(r)$  patterns. Again, bimodal patterns are observed. Interestingly, the separation of both humps is more pronounced for particle 2, indicating that its respective hPhTET3 dimers are, on average, more widely separated than those in particle 1.

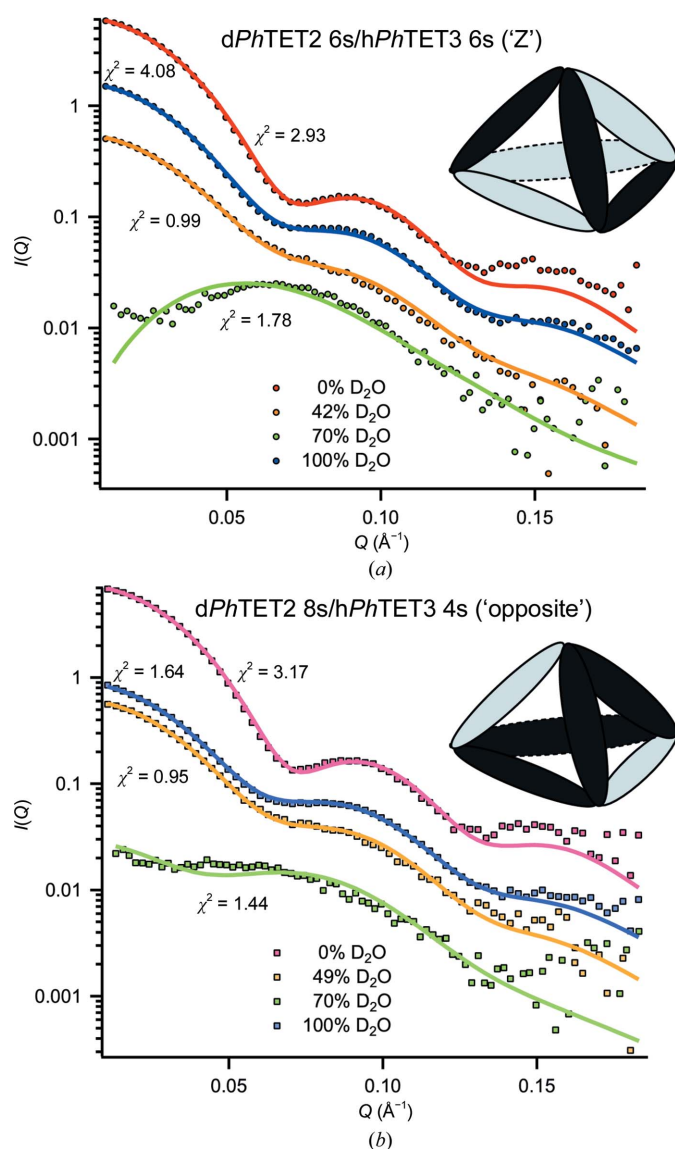
At 70% D<sub>2</sub>O hPhTET3 and dPhTET2 building blocks have negative and positive contrast, respectively (Supplementary

Fig. S4). Since  $p(r)$  functions contain products of pairs of volume elements weighted by their respective contrast, positive and negative regions can be observed in such cases (Glatter & Kratky, 1982): positive  $p(r)$  values occur if volume elements separated by the distance  $r$  are predominantly of identical contrast (positive–positive or negative–negative) and negative values are observed if volume elements of opposite contrast (positive–negative or negative–positive) dominate at this distance. The positive peak at short distances ( $r \approx 20$  Å) therefore belongs to volume elements within dPhTET2 (or hPhTET3) dimeric building blocks. At intermediate distances ( $r \approx 60$  Å) volume elements of opposite contrast (dPhTET2–hPhTET3 pairs) prevail, yielding a minimum, while at large distances ( $r \approx 100$ – $110$  Å) pairs of volume elements with the same sign again dominate. Intriguingly, at 70% D<sub>2</sub>O the minimum at 60 Å is more pronounced for particle 1 than particle 2, indicating that in the former the dPhTET2 and hPhTET3 dimers are distributed more symmetrically (in number and geometry).

### 3.3. Pseudo-atomic models of PhTET2–PhTET3 complexes using SANS contrast-variation data

Fig. 1 illustrates the workflow of our approach to determine pseudo-atomic, quaternary structures of the heterododecameric PhTET2–PhTET3 particles. As a first step, homododecameric ('12s') reference structures of dPhTET2 and hPhTET3 were measured by SANS at 42 and 100% D<sub>2</sub>O, respectively. PhTET2 and PhTET3 dodecameric crystal structures (PDB entries 1y0r and 2wzn, respectively) were slightly modified to match these SANS reference data at 'low resolution' (Fig. 1, top) and were subsequently used to construct a library of heterododecameric models of variable stoichiometry and geometry (Figs. 1 and 2, Supplementary Fig. S13). All structures from the library were scored against the SANS data from peak 1 and peak 2 at 0, 42 (49), 70 and 100% D<sub>2</sub>O. A colour code (green, yellow and red) was used to design excellent, moderate and unacceptable fits. Only one architecture was able to satisfy all contrast conditions simultaneously for each peak and all others could readily be rejected (Fig. 2, Supplementary Fig. S13). Peak 1 was only fitted by a dPhTET2 6s/hPhTET3 6s architecture with the two hexamer ('6s') moieties arranged in an intertwined, double Z-shaped, clamp-like form (Fig. 5a). Peak 2 was only fitted by a dPhTET2 8s/hPhTET3 4s model with the PhTET3 moiety composed of two homodimers ('4s'; four monomers) situated at opposite ridges of the dodecameric particle (Fig. 5b).

In addition, and in complement, to the pseudo-atomic models, the multiphase *ab initio* program MONSA (Svergun, 1999) was used to calculate low-resolution envelopes of the PhTET2 and PhTET3 moieties within the heterododecameric particles of peak 1 and peak 2. A comparison of the low-resolution models with the pseudo-atomic models is shown in Supplementary Fig. S7. The *ab initio* shapes confirm, independently of the pseudo-atomic models, that PhTET2 and PhTET3 are organized as a symmetric architecture of two intertwined Z-shapes in particle 1 and that PhTET2 is



**Figure 5**  
SANS data and best models for peak 1 and peak 2: the two architectures ('Z' and 'opposite') that are in best agreement with the SANS data of peak 1 (a) and peak 2 (b), along with their respective fits at four different contrasts. Elongated ellipsoids represent dPhTET2 (dark grey) and hPhTET3 (light grey) homodimers. Intriguingly, each tetrahedral apex brings together catalytic sites of different substrate specificity.



organized as two homodimers positioned on opposite ridges in particle 2.

### 3.4. The homododecameric *PhTET2* complex is extremely stable in solution and is not in dynamic equilibrium with free oligomeric forms of lower molecular mass

As an additional stability control of the *PhTET* quaternary structure and to test the presence of putative equilibria between preformed dodecameric particles with smaller oligomeric particles in solution, we measured mixtures of pre-formed homododecameric deuterated and hydrogenated *PhTET2* proteins (d*PhTET2* '12s' and h*PhTET2* '12s') by monitoring their SANS curves as a function of temperature and exposure time. The experiments were performed on both the wild-type protein and a pentamutant (R217S, R220S, F224S, H248S and I292A) with the mutations situated at the interface between the dimers and designed to slow the oligomerization process (Appolaire *et al.*, 2013). For the wild-type and the mutated *PhTET2*, the following samples were prepared: h*PhTET2* (42% D<sub>2</sub>O), d*PhTET2* (42% D<sub>2</sub>O), h*PhTET2* (100% D<sub>2</sub>O) and d*PhTET2* (100% D<sub>2</sub>O). Different mixtures of the above samples were prepared by mixing 1:1 volume fractions of these solutions.

The SANS curves of isolated h*PhTET2* (100% D<sub>2</sub>O) and d*PhTET2* (42% D<sub>2</sub>O) reference particles (Supplementary Fig. S3) revealed the typical pattern of the dodecameric edifice as

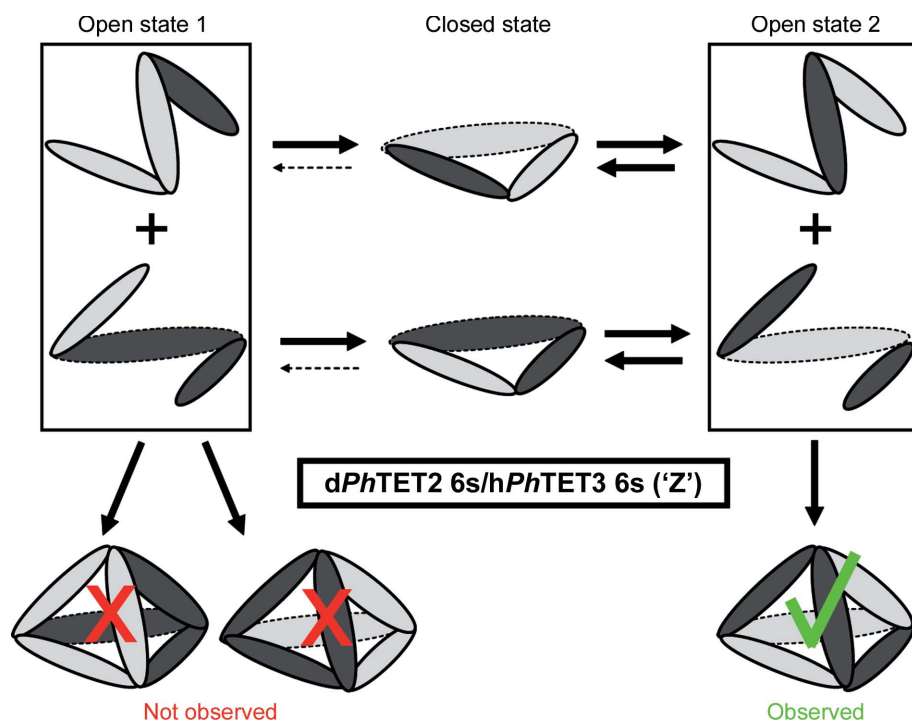
observed in previous SANS experiments (Durá *et al.*, 2009). Supplementary Fig. S3(a) shows a superposition of the scattering curves of wild-type h*PhTET2* in 100% D<sub>2</sub>O at  $t = 0$  (20°C) and of a 1:1 mixture of d*PhTET2* and h*PhTET2* in 100% D<sub>2</sub>O after 4 h at 80°C, both normalized to h*PhTET2* concentration. Supplementary Fig. S3(b) shows the respective data sets of a 1:1 d*PhTET2*/h*PhTET2* mixture at 42% D<sub>2</sub>O. Interestingly, the scattering patterns of both mixtures after 4 h at 80°C are identical to those of their respective dodecameric *PhTET2* references at 20°C regarding radii of gyration, position of the side maxima and minima and  $I(0)$  intensities. These very clear results indicate that no exchange of lower oligomeric building blocks between the hydrogenated and deuterated dodecamers had occurred in solution on the temperature and time scales of our SANS experiments. These experiments confirm the great stability of both the wild-type and pentamutant dodecameric particles once they are formed in solution.

## 4. Discussion

Large and symmetric complexes of multiple copies of a single or a few similar building blocks play essential roles in the life cycle of biological cells (Matthews & Sunde, 2012; Griffin & Gerrard, 2012; Sauer & Baker, 2011; Snider & Houry, 2008; Marianayagam *et al.*, 2004). If crystal structures are not available, they represent a formidable challenge for many

structural biology techniques: for NMR owing to their size and for SAXS or EM because it is very difficult to position individual subunits within the assembled complexes. The heterododecameric *PhTET2*–*PhTET3* complexes studied here are highly symmetric and we applied a powerful combination of SANS, deuterium labelling and contrast variation to obtain unique insights into the oligomeric organization of their quaternary architectures. In contrast to EM or SAXS, the neutron scattering lengths for hydrogen and deuterium differ significantly and are of opposite sign (Jacrot, 1976). SANS can therefore focus specifically on D- or H-labelled partners within a reconstituted complex. Using this approach, we were able to eliminate a multitude of potential stoichiometries and geometric arrangements very efficiently and to clearly identify a single, specific architecture of the *PhTET2* and *PhTET3* moieties within the final heterododecameric complexes from peak 1 and peak 2 (Fig. 2, Supplementary Fig. S13).

Previously, a combination of site-directed mutagenesis, SAXS, AUC and



**Figure 6**

Assembly pathway for the 'Z' d*PhTET2* 6s:h*PhTET3* 6s heterododecamer (peak 1): the observation of a specific internal hexamer–hexamer topology by SANS, together with the exclusion of alternative topologies, points to a specific assembly pathway involving intertwined heterohexamers. Thick continuous and thin broken arrows represent strong and weak pathways for the dynamic equilibrium between respective states. An alternative assembly mode is presented in Supplementary Fig. S8.

EM allowed us to propose an assembly mechanism in which three dimers associate into a hexameric precursor prior to the formation of TET dodecamers (Appolaire *et al.*, 2013). This model accounts well for the very limited number of hetero-oligomeric TET complexes that we have identified in the present work by ion-exchange chromatography after metal-induced re-oligomerization of a mixture of *PhTET2* and *PhTET3* dimers (Fig. 3). This strategy was possible because the *PhTET2* and *PhTET3* partners behave differently on strong ion-exchange chromatography, allowing the separation of different hetero-oligomeric complexes as a function of their *PhTET2:PhTET3* ratio. The SANS study described here provides direct insight into this assembly model. Indeed, the topology of the *PhTET2* and *PhTET3* moieties within the two main hetero-oligomeric species observed are in excellent agreement with an assembly process that involves a hexameric intermediate (Fig. 6, Supplementary Fig. S8), as suggested by a previous study combining SAXS, EM and AUC (Appolaire *et al.*, 2013).

Each of the TET apices defines a catalytic subcompartment in which three active sites are located on the same plane. Moreover, a nonprocessive mode of action has been reported for the three TET proteins from *P. horikoshii* (Durá *et al.*, 2005, 2009; Schoehn *et al.*, 2006). Therefore, after the cleavage of the N-terminal amino acid the polypeptide substrate is released from the active site and is free to interact with one of the two other catalytic pockets. In the present study we showed that, unexpectedly, TET dodecameric quaternary structures follow a highly organized pathway when self-assembling *in vitro*. This results in a limited number of architectures with a single type of apex composition, systematically combining *PhTET2* and *PhTET3* active sites with different and complementary substrate specificities (Fig. 6), suggesting that the highly controlled oligomerization process of the *PhTET2-PhTET3* complex is to optimize its peptide-degradation efficiency.

Finally, our work shows that SANS experiments allow the respective positions of the different subunits to be very efficiently specified in the case of a large symmetrical pseudo-homomeric complex, information that is very difficult, if not impossible, to obtain using most biophysical techniques. In conclusion, such a SANS approach could be very efficient to better understand the assembly pathway of other large, symmetric, pseudo-homomeric complexes such as the proteasome (Sahara *et al.*, 2014), chaperonins (Kim *et al.*, 2013) or AAA-ATPases (Bar-Nun & Glickman, 2012). Our structural insights raise intriguing questions on the underlying regulatory mechanisms controlling the assembly process of such biomacromolecular complexes, a field of research that has remained poorly explored to date but might offer important perspectives for biomedical research and drug development in the future.

We acknowledge financial support from the Agence Nationale de la Recherche (grants 'MacroTET' BLAN-07-3204002, 'Archelyse' ANR-12-BSV8-0019-01 and 'HYDROSAS' ANR-11-JSV5-003-01). AA was supported by

a PhD scholarship from the French Ministry for Research and Technology. MAD and MC were supported by a French National Research Agency postdoctoral fellowship. We would like to thank the ILL for SANS BAG beamtime on D22 and P. Callow and A. Martel for help with the set-up of the SANS instrument. MM and MH acknowledge EPSRC support (EP/C015452/1) to V. T. Forsyth (Keele University, England) for the creation of the Deuteration Laboratory within ILL's Life Sciences Group and from the EU under contract RII3-CT-2003-505925.

## References

- Appolaire, A., Rosenbaum, E., Durá, M. A., Colombo, M., Marty, V., Savoye, M. N., Godfroy, A., Schoehn, G., Girard, E., Gabel, F. & Franzetti, B. (2013). *J. Biol. Chem.* **288**, 22542–22554.
- Artero, J.-B., Härtlein, M., McSweeney, S. & Timmins, P. (2005). *Acta Cryst.* **D61**, 1541–1549.
- Bar-Nun, S. & Glickman, M. H. (2012). *Biochim. Biophys. Acta*, **1823**, 67–82.
- Borissenko, L. & Groll, M. (2005). *J. Mol. Biol.* **346**, 1207–1219.
- Chaikuad, A., Pilka, E. S., De Riso, A., von Delft, F., Kavanagh, K. L., Vénien-Bryan, C., Oppermann, U. & Yue, W. W. (2012). *BMC Struct. Biol.* **12**, 14.
- Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K. & Kiser, P. D. (2012). *J. Biol. Chem.* **287**, 13356–13370.
- Durá, M. A., Receveur-Brechot, V., Andrieu, J.-P., Ebel, C., Schoehn, G., Roussel, A. & Franzetti, B. (2005). *Biochemistry*, **44**, 3477–3486.
- Durá, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M. & Franzetti, B. (2009). *Mol. Microbiol.* **72**, 26–40.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Franzetti, B., Schoehn, G., Hernandez, J. F., Jaquinod, M., Ruigrok, R. W. & Zaccari, G. (2002). *EMBO J.* **21**, 2132–2138.
- Glatter, O. & Kratky, O. (1982). *Small Angle X-ray Scattering*. New York: Academic Press.
- Gosh, R. E., Egelhaaf, S. U. & Rennie, A. R. (2006). *A Computing Guide for Small-angle Scattering Experiments*. Grenoble: Institut Laue-Langevin.
- Griffin, M. & Gerrard, J. (2012). *Protein Dimerisation and Oligomerisation in Biology*, edited by J. Matthews, pp. 74–90. New York: Springer.
- Guinier, A. (1939). *Ann. Phys. (Paris)*, **12**, 161–236.
- Jacrot, B. (1976). *Rep. Prog. Phys.* **39**, 911–953.
- Jacrot, B. & Zaccari, G. (1981). *Biopolymers*, **20**, 2413–2426.
- Kim, D., San, B. H., Moh, S. H., Park, H., Kim, D. Y., Lee, S. & Kim, K. K. (2010). *Biochem. Biophys. Res. Commun.* **391**, 431–436.
- Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. & Hartl, F. U. (2013). *Annu. Rev. Biochem.* **82**, 323–355.
- Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). *Q. Rev. Biophys.* **36**, 147–227.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
- Marianayagam, N. J., Sunde, M. & Matthews, J. M. (2004). *Trends Biochem. Sci.* **29**, 618–625.
- Matthews, J. & Sunde, M. (2012). *Protein Dimerisation and Oligomerisation in Biology*, edited by J. Matthews, pp. 1–18. New York: Springer.
- Rawlings, N. D., Waller, M., Barrett, A. J. & Bateman, A. (2014). *Nucleic Acids Res.* **42**, D503–D509.
- Rosenbaum, E., Ferruit, M., Durá, M. A. & Franzetti, B. (2011). *Biochim. Biophys. Acta*, **1814**, 1289–1294.
- Rosenbaum, E., Gabel, F., Durá, M. A., Finet, S., Cléry-Barraud, C., Masson, P. & Franzetti, B. (2012). *Arch. Biochem. Biophys.* **517**, 104–110.
- Russo, S. & Baumann, U. (2004). *J. Biol. Chem.* **279**, 51275–51281.

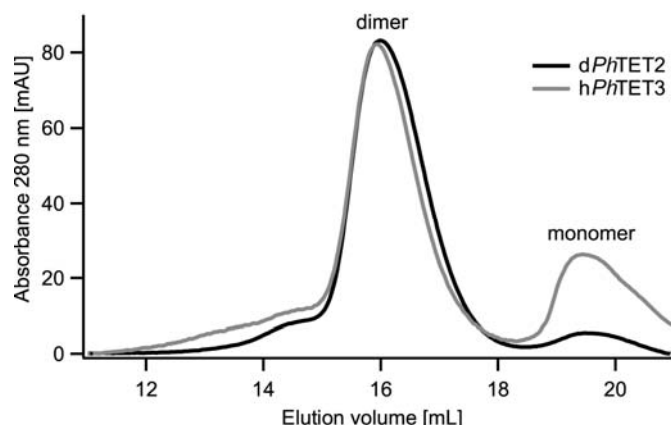
- Sahara, K., Kogleck, L., Yashiroda, H. & Murata, S. (2014). *Adv. Biol. Regul.* **54**, 51–58.
- Sauer, R. T. & Baker, T. A. (2011). *Annu. Rev. Biochem.* **80**, 587–612.
- Schoehn, G., Vellieux, F. M. D., Durá, M. A., Receveur-Bréchet, V., Fabry, C. M. S., Ruigrok, R. W. H., Ebel, C., Roussel, A. & Franzetti, B. (2006). *J. Biol. Chem.* **281**, 36327–36337.
- Snider, J. & Houry, W. A. (2008). *Biochem. Soc. Trans.* **36**, 72–77.
- Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
- Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.
- Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
- Svergun, D. I. & Koch, M. H. J. (2002). *Rep. Prog. Phys.* **66**, 1735–1782.
- Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, Z., Kuprin, S. & Zaccai, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2267–2272.



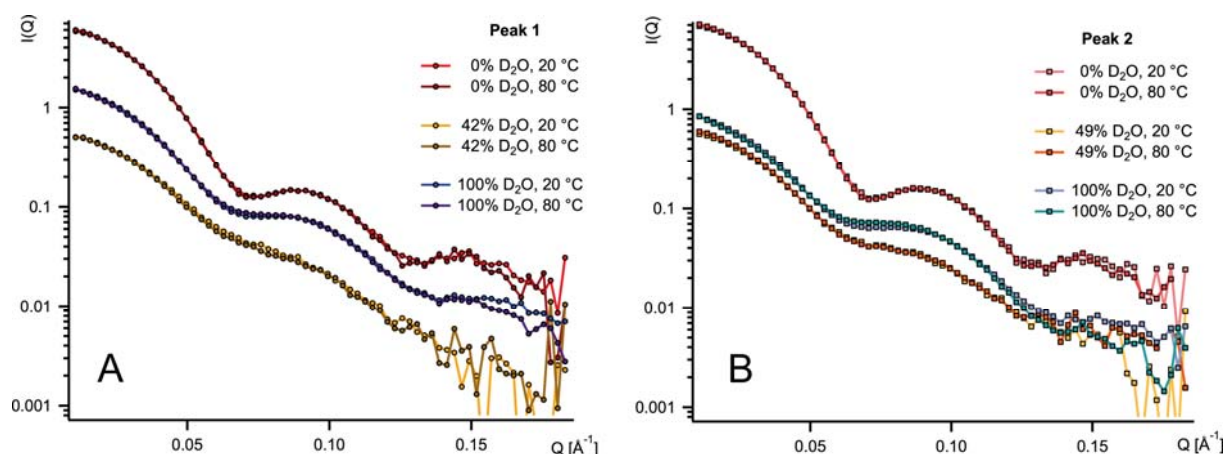
# Supplementary information

## Small angle neutron scattering reveals the assembling mode and oligomeric organization of TET, a large dodecameric aminopeptidase

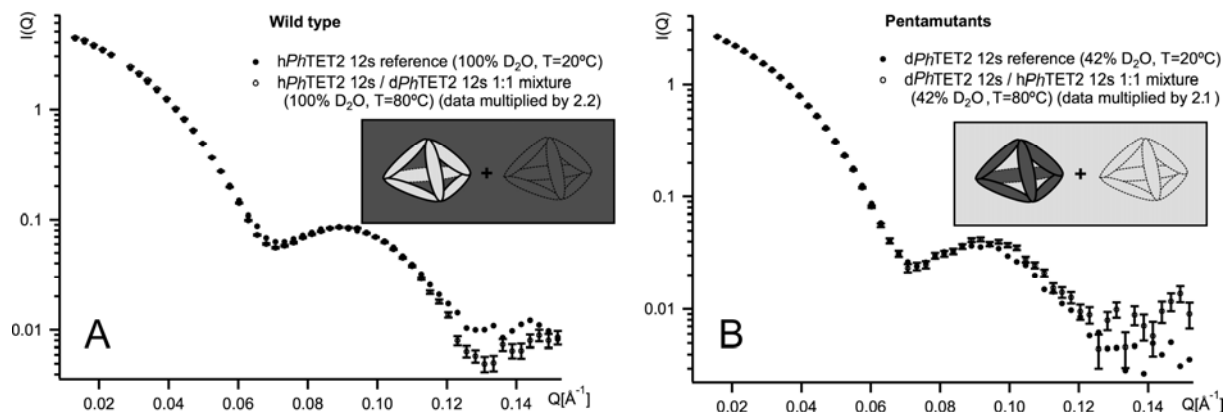
Alexandre Appolaire, Eric Girard, Matteo Colombo, M. Asunción Durá, Martine Moulin, Michael Härtlein, Bruno Franzetti and Frank Gabel



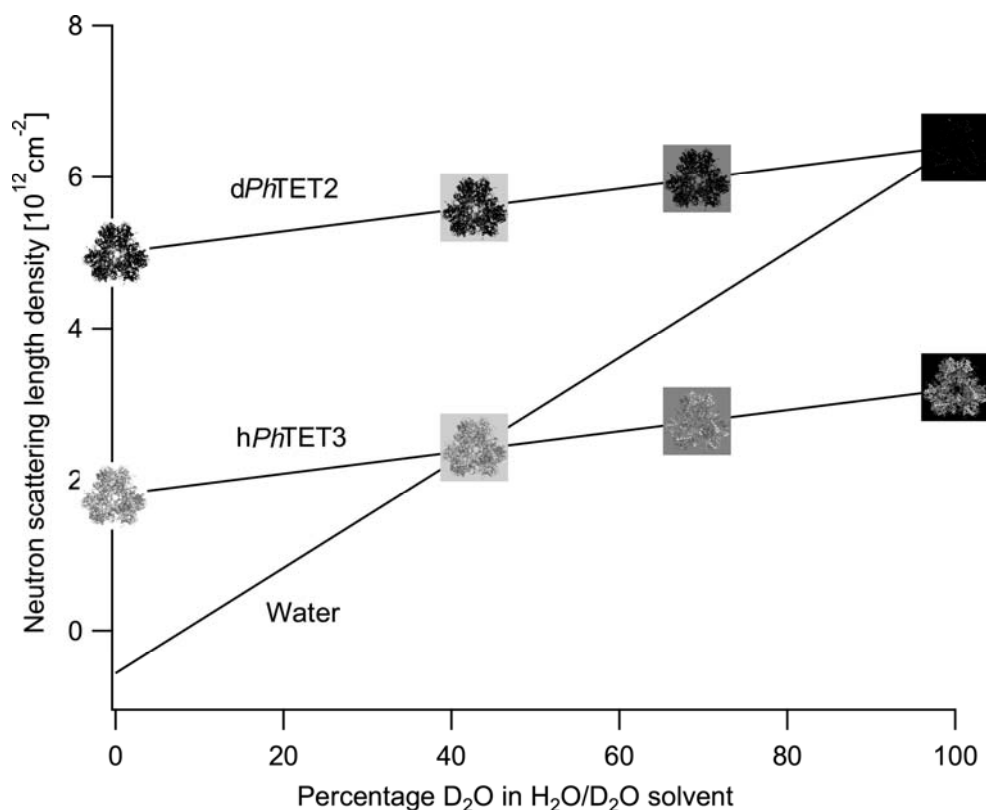
**Fig. S1: Elution profile on a size exclusion column (Superose 6) of the dPhTET2 and hPhTET3 samples immediately after the de-oligomerization step.** hPhTET3 contains approximately 30% of monomers, while dPhTET2 yielded less than 4% of monomers, limiting the overall probability to find hetero-dimers to a maximum of 4%.



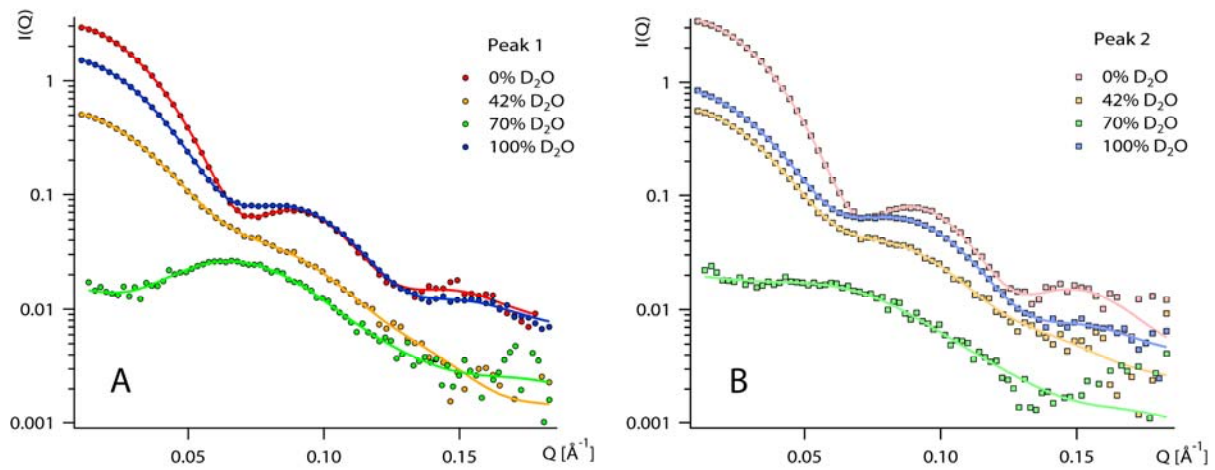
**Fig. S2: Temperature dependence of the hetero-dodecamer SANS curves.** Experimental SANS data of Peak 1 (A) and Peak 2 (B) at three contrast conditions at 20°C and after four hours at 80°C. No major changes in shape or intensity were observed between the respective datasets, indicating that both the oligomeric state and the shape of the two hetero-dodecameric complexes are stable over several hours at 80°C. A slight flattening of the first minimum is observed at 80°C, possibly due to an increased conformational flexibility of loops, C- and N-terminal parts at higher temperatures. 70% D<sub>2</sub>O datasets were not recorded at 80 °C.



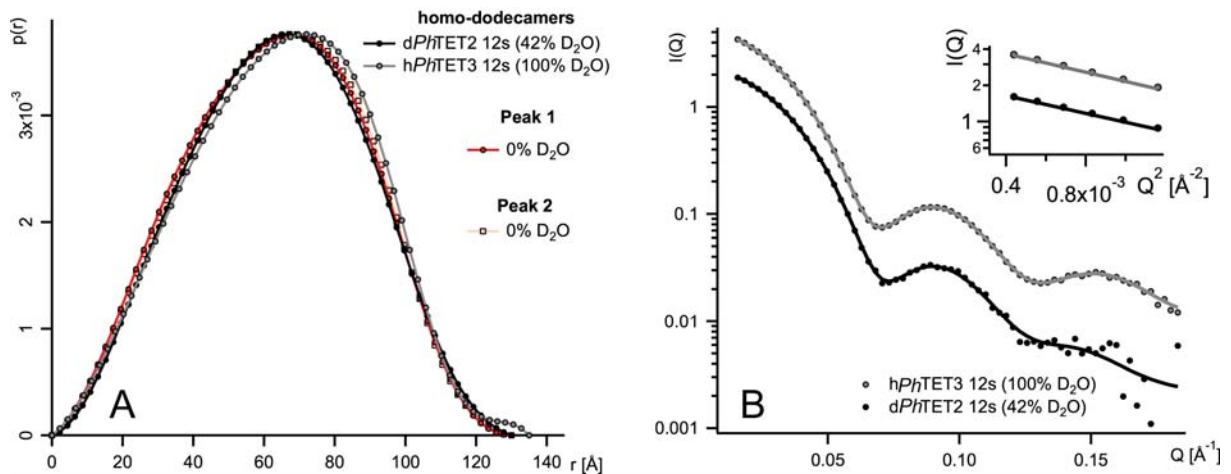
**Fig S3. Stability of *hPhTET2* and *dPhTET2* homo-dodecamers (“12s”).** A) Superposition of SANS curves from dodecameric wild type *hPhTET2* alone in solution at 20°C (reference) and a 1:1 mixture of *hPhTET2* and *dPhTET2* at 80°C, both in 100% D<sub>2</sub>O. The signal of the mixture was scaled by a factor of 2.2 which corresponds well to the factor expected from the dilution. Pentamutant *hPhTET2* samples displayed an equivalent behavior. B) Superposition of SANS curves from pentamutant dodecameric *dPhTET2* alone in solution at 20°C (reference) and a 1:1 mixture of *hPhTET2* and *dPhTET2* pentamutants at 80°C, both in 42% D<sub>2</sub>O. The signal of the mixture was scaled by a factor of 2.1 which agrees well with the factor expected from the dilution. Wild-type *dPhTET2* samples displayed an equivalent behavior.



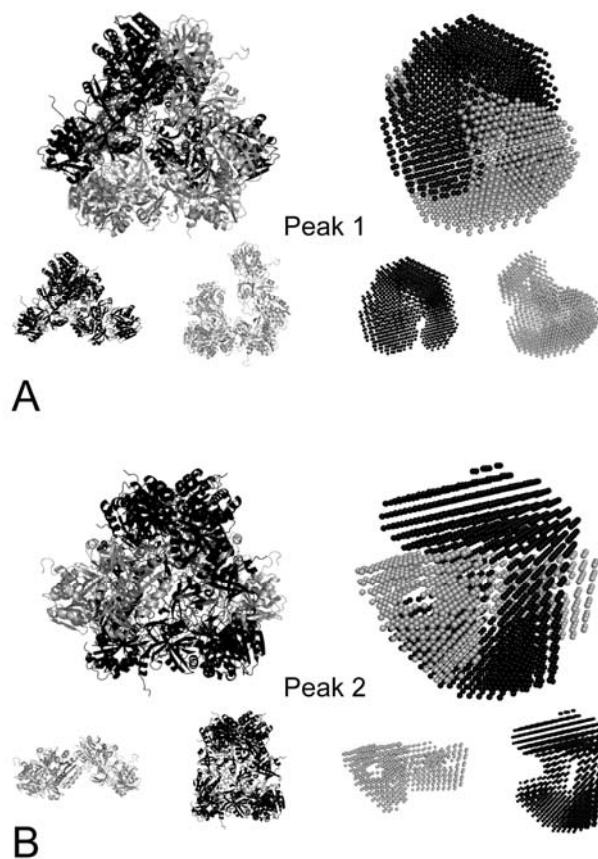
**Fig. S4: Representation of the contrast conditions for *dPhTET2* and *hPhTET3*.** Neutron scattering length densities of solvent, *dPhTET2* and *hPhTET3*. *dPhTET2* (black) and *hPhTET3* (gray) are depicted in front of a white, gray, dark gray or black background representing 0, 42, 70 and 100% D<sub>2</sub>O in the solvent, respectively (This color code is applied throughout the present work). *PhTET* particles that have positive contrast are above the water solvent line, those with negative contrast are below.



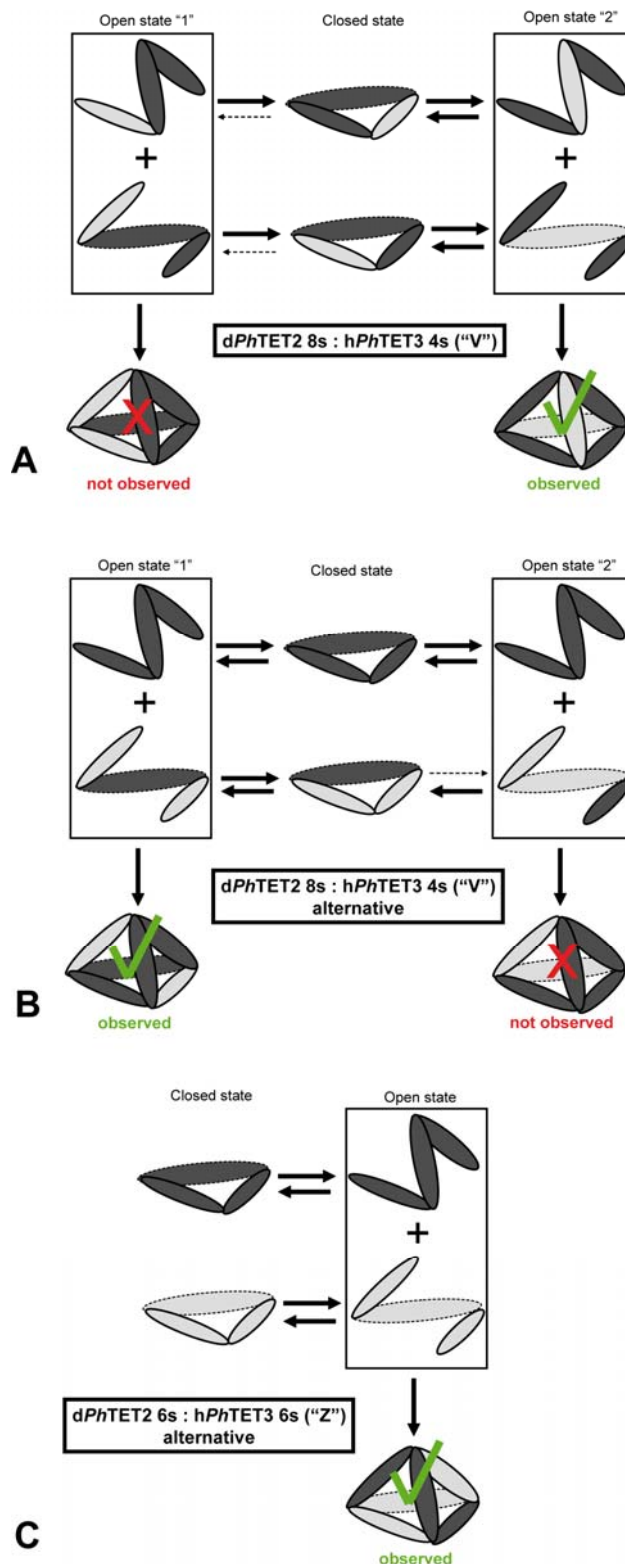
**Fig. S5: GNOM fits of the SANS data at 20°C.** Experimental SANS data of Peak 1 (A) and Peak 2 (B) at the four contrast conditions (filled symbols). The continuous lines of identical colors show the respective regularized curves calculated by the program GNOM. These regularizations were used to extract the  $p(r)$  functions shown in Fig. 2. In all cases (apart from the 70% D<sub>2</sub>O datasets) the fits were judged as “excellent” by GNOM (“reasonable” for the 70% D<sub>2</sub>O data).



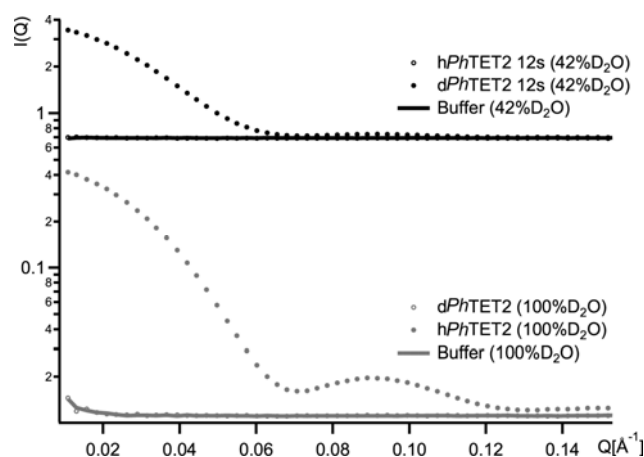
**Fig. S6:  $p(r)$  functions of hetero- and homo-oligomers.** A) Comparison of pair-distance distribution functions of the homo-dodecameric particles with respect to the hetero-dodecamers. B) GNOM back-calculated scattering curves against the respective experimental SANS curves for the homo-dodecameric particles. The respective Guinier fits are shown as an inset. The agreement is excellent. All datasets shown were recorded at 20°C.



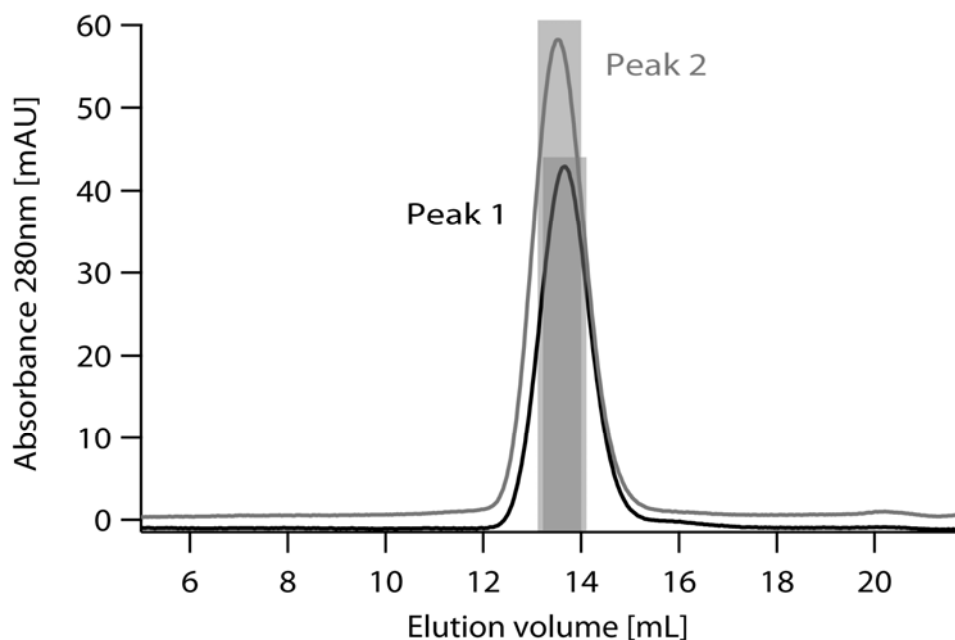
**Fig. S7: A) Comparison of the final quasi-atomic model of Peak 1 and its low-resolution reconstitution from MONSA.** The individual dPhTET2 (black) and hPhTET3 (gray) moieties are shown as an inset at the bottom. **B) Comparison of the final quasi-atomic model of Peak 2 and its low-resolution reconstitution from MONSA.** The individual dPhTET2 (black) and hPhTET3 (gray) moieties are shown as an inset at the bottom.



**Fig. S8: Assembling pathways for the “opposite” 8s4s (Peak 2) and the “Z” 6s6s (Peak 1) hetero-dodecamers:** A) and B) all possible assembling modes for a hetero-complex consisting of four *dPhTET2* and two *hPhTET3* dimers based on intermediate hexamers. The observed forms (two opposite *hPhTET3* dimers) display mixed *PhTET2:PhTET3* apices and therefore heterogeneous catalytic chambers at each apex. C) An alternative pathway for the “Z” 6s6s particle from Fig. 5 which yields the same final architecture. In all cases, bold, continuous and broken fine arrows represent strong and weak pathways for the dynamic equilibrium between respective states.

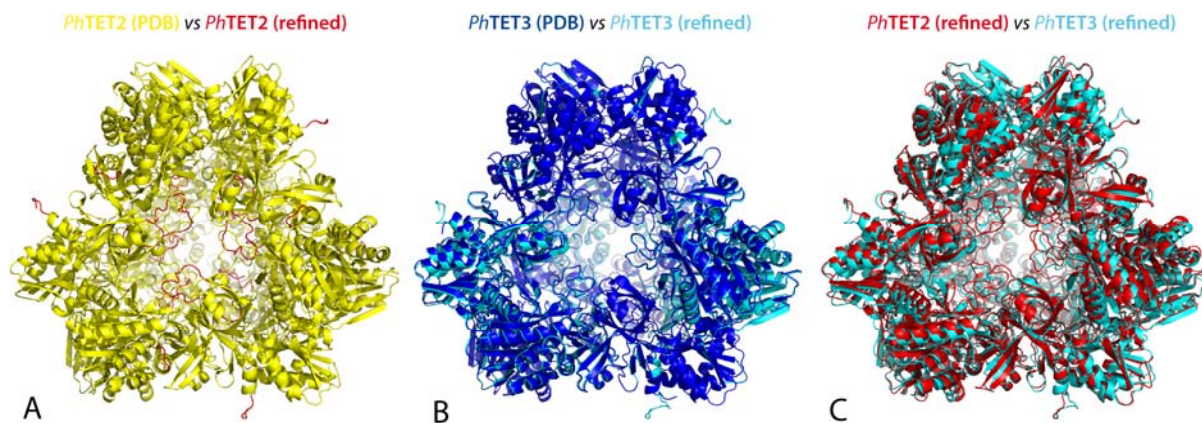


**Fig. S9: Experimental validation of SANS contrast matching conditions at 20°C.** SANS curves of mono-dodecameric *hPhTET2* and *dPhTET2* (wild-type) in 42 and 100% D<sub>2</sub>O (prior to buffer subtraction) and the respective buffers. The upper part of the figure illustrates clearly that the hydrogenated partner (*hPhTET2*) is perfectly matched at 42% D<sub>2</sub>O and that the signal at this contrast point is dominated by the deuterated partner (*dPhTET2*). The lower part demonstrates that the inverse situation is true at 100% D<sub>2</sub>O (the latter data have been shifted for clarity).

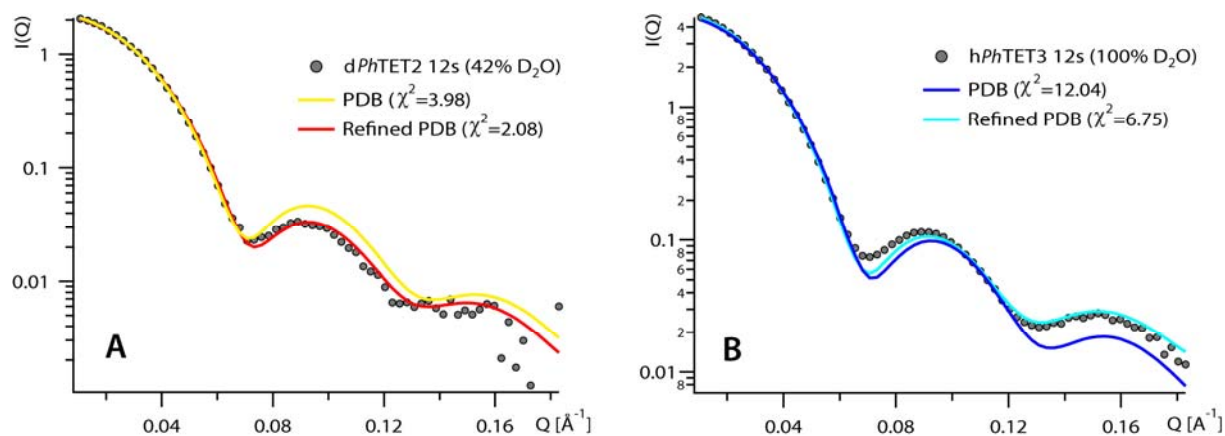


**Fig. S10: Elution profile on a size-exclusion column (Superose 6) of the Mono Q purified peaks 1 and 2.** The fractions shaded in gray were grouped together to constitute the samples measured by SANS.





**Fig. S11: Comparison of *PhTET2* and *PhTET3* homo-dodecameric reference particles for SANS.** A: Original *PhTET2* crystal structure (PDB entry 2Y0R) (yellow). The N-terminus (residues 1-5) and the inner loop (residues 120-132) that have been added to optimize the fit against the SANS reference data (d*PhTET2* 12s in 42% D<sub>2</sub>O) are shown in red. B: Original *PhTET3* crystal structure (PDB entry 2WZN) (blue). The model refined by rigid-body motions to optimize the fit against the SANS reference data (h*PhTET3* 12s in 100% D<sub>2</sub>O) is shown in cyan. C: Superposition of the refined *PhTET2* (red) and *PhTET3* (cyan) models.



**Fig. S12: Comparison of fits of *PhTET2* and *PhTET3* homo-dodecameric reference particles against SANS data (same colour code as in Fig. S11).** A: Original *PhTET2* crystal structure (PDB entry 2Y0R) (yellow) and refined model (red) against the SANS reference data (d*PhTET2* 12s in 42% D<sub>2</sub>O). B: Original *PhTET3* crystal structure (PDB entry 2WZN) (blue) and the refined model (cyan) against the SANS reference data (h*PhTET3* 12s in 100% D<sub>2</sub>O).

	0% D <sub>2</sub> O	42% D <sub>2</sub> O	49% D <sub>2</sub> O	100% D <sub>2</sub> O
hPhTET3 12s	1.06	0.0023	0.066	<b>4.19</b>
hPhTET3 10s / dPhTET2 2s	1.66	0.06	0.0016	2.90
hPhTET3 8s / dPhTET2 4s	2.39	0.30	0.11	1.84
hPhTET3 6s / dPhTET2 6s	<b>3.26</b>	<b>0.71</b>	0.40	<b>1.02</b>
hPhTET3 4s / dPhTET2 8s	<b>4.26</b>	1.30	<b>0.86</b>	<b>0.47</b>
hPhTET3 2s / dPhTET2 10s	5.39	2.07	1.50	0.10
dPhTET2 12s	6.65	<b>3.01</b>	2.31	0.001
<b>"Peak1" (experimental)</b>	<b>3.28</b>	<b>0.56</b>	-	<b>1.66</b>
<b>"Peak2" (experimental)</b>	<b>3.84</b>	-	<b>0.62</b>	<b>0.90</b>
<b>hPhTET3 12s (experimental)</b>	-	-	-	<b>5.28</b>
<b>dPhTET2 12s (experimental)</b>	-	<b>2.32</b>	-	-

**Table S1: Comparison of measured with calculated  $I(0)$  intensities.** The measured  $I(0)$  intensities are the same as the ones in Table 1 of the main document. The theoretical ones (first seven data lines) were calculated according to (Jacrot & Zaccai 1981), using the following

equation:

$$I(0) = f \frac{4\pi T_s}{1-T} I_{inc}(0) \frac{C}{M_r} N_A t 10^{-3} \left( \sum_i b_i - \rho_s V \right)^2 \quad (\text{Suppl. Eq. 1})$$

$f=0.8$  is correction factor taking the anisotropic scattering of the incoherent background into account,  $T_s$  and  $T$  are the sample and H<sub>2</sub>O transmissions, respectively,  $I_{inc}(0)$  is the incoherent water background at zero angle,  $C=4.5$  mg/ml are the sample concentrations,  $N_A$  is Avogadro's constant,  $t=0.1$ cm is the sample cell pathlength.  $\sum_i b_i$  and  $V$  are summed protein

scattering lengths and solvent-excluded volumes, calculated from their amino acid sequences (Jacrot 1976) and  $\rho_s$  is the solvent scattering length density at the respective contrast conditions.  $\sum_i b_i$  were calculated assuming that all exchangeable hydrogens exchange. The

precision of the theoretical intensities is limited by the precision of the sample concentrations (determined by Bradford in our case) but depends also strongly on solvent-excluded volumes and the fraction of exchangeable hydrogens that is assumed. We estimate the overall precision with  $\pm 30\%$ . As can be seen from the two reference data sets (hPhTET3 12s and dPhTET2 12s), the discrepancy between experimental and theoretical values can be both positive or negative.

### 6.3 DISCUSSION ET PERSPECTIVES :

---

Le travail présenté ici démontre le caractère non aléatoire de l'assemblage d'un hétéro-oligomère de la peptidase TET. Cette étude a été possible grâce à l'utilisation de la diffusion de neutrons aux petits angles (SANS) et en particulier à la technique de variation de contraste. D'un point de vue méthodologique, ce travail démontre le grand intérêt que représente cette technique puisqu'elle a permis de révéler la topologie d'un assemblage de grande symétrie sans ambiguïté et de proposer une voie d'oligomérisation. La réalisation de ces expériences n'a été possible qu'avec le travail de biochimie réalisé en amont. En effet, sans les connaissances liées aux propriétés d'assemblage de la particule TET, il aurait été difficile de pouvoir produire un hétéro-oligomère dont seul une partie des sous-unités serait deutérée. De plus, les propriétés de charges différentes des sous-unités ont permis un criblage des protéines d'intérêt en amont de l'expérience, notamment grâce à l'utilisation de la chromatographie d'affinité. Enfin, le fait que les structures cristallographiques soient disponibles a permis de créer un modèle sur lequel s'appuyer pour analyser les données expérimentales. C'est donc une technique extrêmement puissante pour l'étude de grands assemblages en solution, mais qui nécessite de s'inscrire dans une stratégie intégrative et non pas comme seule technique d'investigation.

D'autre part, d'un point de vue biochimique, il est intéressant de noter que, parmi les différentes possibilités d'assemblage, les complexes prépondérants observés présentent une topologie ne comprenant que des poches catalytiques constituées au moins de deux sous-unités différentes. L'aminopeptidase TET étant non processive, et chaque poche catalytique composée de trois sites actifs, cette propriété lui permettrait d'hydrolyser plusieurs résidus de natures différentes sans que le peptide substrat n'ait quitté la poche catalytique. De plus, ce type de topologie n'est possible que par un nombre limité de voies d'assemblage. Le modèle d'oligomérisation présenté dans le chapitre précédent propose que l'intermédiaire principal d'oligomérisation soit un hexamère de TET.

La composition des intermédiaires hexamériques impliqués dans les voies d'oligomérisation possibles suggère une affinité accrue entre PhTET2 et PhTET3. En effet, le mode d'assemblage proposé montre que dans un hexamère formé de deux sous-unités PhTET2 et une sous-unité PhTET3, le point d'ouverture de l'hexamère menant à

la formation du dodécamère se ferait au niveau de l'interface entre les deux sous-unités PhTET2, et *vice versa*. L'ensemble de ces résultats, ainsi que ceux obtenus dans le chapitre 4, conforte l'idée qu'*in vivo*, la formation d'hétéro-oligomères de PhTET2 et PhTET3 serait favorisée et que cette particule représenterait un peptidasome dans la cellule de *Pyrococcus horikoshii*.



## 7. RESULTATS PRELIMINAIRES DE LA CARACTERISATION DE PHTET4

L'étude de la phylogénie des TET présentée dans le chapitre 3 a permis de mettre en évidence l'existence d'une quatrième TET chez *P.horikoshii*. J'ai pu débiter sa caractérisation durant ma thèse, ce chapitre regroupe les résultats préliminaires obtenus.

### 7.1 IDENTIFICATION

---

La phase ouverte de lecture PH0737 de *Pyrococcus horikoshii* a été identifiée comme codant pour une quatrième TET. PH0737 code pour une protéine de 336 acides aminés dont le poids moléculaire estimé est de 37 kDa. La séquence de la protéine identifiée montre un pourcentage de similarité de 23 % (45 % d'homologie) pour PhTET1 et de 25 % (44 % d'homologie) pour PhTET2 et PhTET3. Parmi le groupe de quatre TET de *P.horikoshii*, celle-ci paraît donc être la plus éloignée. En effet, PhTET1 montre environ 37 % d'identité (55 % d'homologie) pour PhTET2 et PhTET3, et ces deux dernières ont 48 % d'identité (70 % d'homologie) entre elles. Elle est annotée comme étant une peptidase de la famille M42 homologue de la peptidase putative yhfE de *Bacillus subtilis*, ce dernier point est cohérent avec les données obtenues lors de l'étude phylogénétique montrant l'origine bactérienne des TET de *P. horikoshii*.

L'alignement des séquences des quatre TET présenté en Figure 7-1 montre une conservation du site actif, ce qui explique son assignation en tant que M42. Les critères pour l'assignation d'une peptidase en tant que TET ne sont pas tous réunis. Le domaine dimérisation est bien présent mais on remarque que le motif « IDIGAXXXE » n'est pas parfaitement conservé. De plus le domaine correspondant à l'hélice  $\alpha 5$  semble être également différent. Enfin les glycines identifiées comme importantes pour le repliement du domaine catalytique sont conservées.



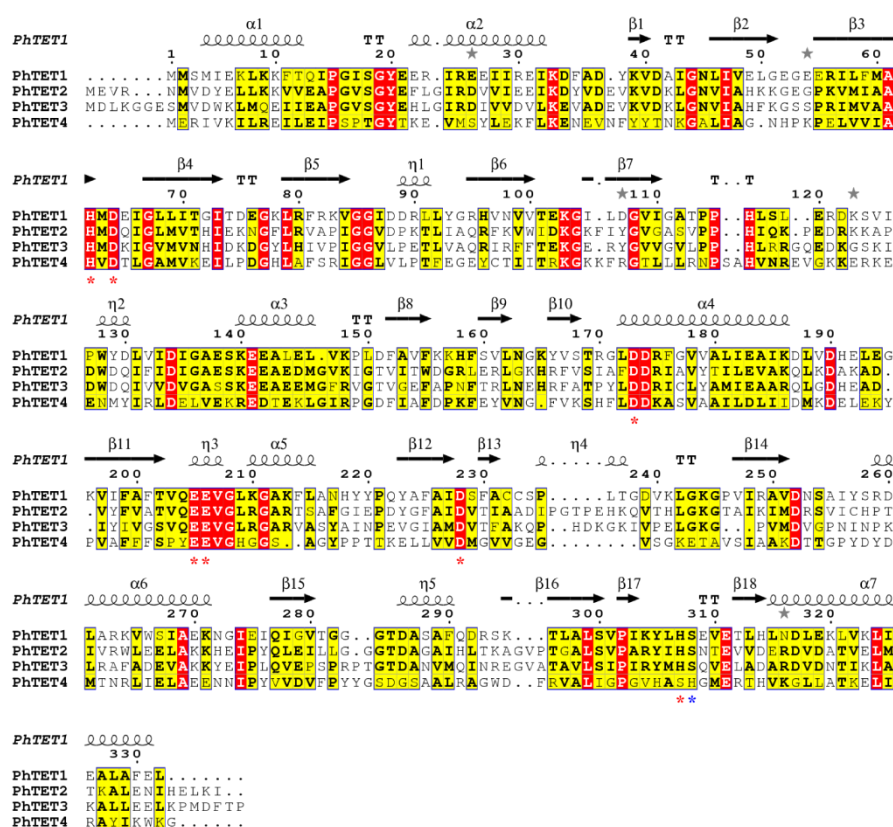


Figure 7-1 : Alignement multiple des séquences des TET de *Pyrococcus horikoshii* réalisé avec clustalW2. Les structures secondaires indiquées correspondent à celles de PhTET1. Les astérisques rouges indiquent les résidus conservés du site actif. L'astérisque bleu en position 308 indique la position de l'histidine du site actif chez PhTET4.

### 7.2 PURIFICATION

Le gène de PhTET4 a été inséré dans un plasmide pET41c et transformé dans des cellules de la souche d'*Escherichia coli* BL21C+ DE3 RIL. Le protocole de transformation et d'expression est le même que celui utilisé pour PhTET2 présenté dans le chapitre 5.

La mise au point des conditions de purification s'est déroulée en plusieurs étapes. Dans un premier temps, le protocole « classique » de purification des TET a été utilisé tel que décrit dans la section 2.1.1. Cependant, la protéine PhTET4 présentait une stabilité apparemment moins élevée que les autres PhTET. Les premiers essais de purification ont montré un rendement quasi nul. Plusieurs séries d'essai de purification ont permis d'identifier les deux étapes durant lesquelles les protéines précipitaient : ce protocole compte une étape de choc thermique à 85 °C durant laquelle près de 50 % de la protéine précipitait, le reste précipitait durant la dialyse précédant la première colonne échangeuse d'ion.

Ceci m'a conduit à tester plusieurs autres méthodes de purification afin d'améliorer le rendement de purification de la protéine. Le but du choc thermique de ce protocole est

d'éliminer les protéines de l'organisme hôte, j'ai tenté de le remplacer par une étape de gradient de sucrose, décrit dans la section 2.1.1, qui n'a pas résolu le problème, le rendement dans ce cas étant quasi nul. L'essentiel des protéines d'intérêt a précipité durant la centrifugation.

Durant les dernières purifications réalisées dans le cadre de la thèse, un protocole comptant une étape de choc thermique à 75 °C et une étape de dilution précédant la colonne échangeuse d'ions plutôt qu'une dialyse a été utilisé. A la suite de ces étapes, l'échantillon a été analysé sur colonne échangeuse d'ions ResourceQ équilibré avec une solution à 100 mM NaCl, 20 mM Tris-HCl, pH 7,5. Les protéines adsorbées ont été éluées en appliquant un gradient linéaire de sel (0.1– 0.35 M NaCl dans 20 mM Tris-HCl, pH 7.5). Les fractions contenant la protéine d'intérêt ont été identifiées par gel SDS-PAGE (Figure 7-2A). Enfin, le pool de protéine a été analysé sur colonne de gel filtration superose 6, un pic majoritaire éluant à environ 14 mL est observé (Figure 7-2B). Ce volume d'élution sur cette colonne est typique de l'assemblage dodécamérique d'une TET (Figure 7-2B).

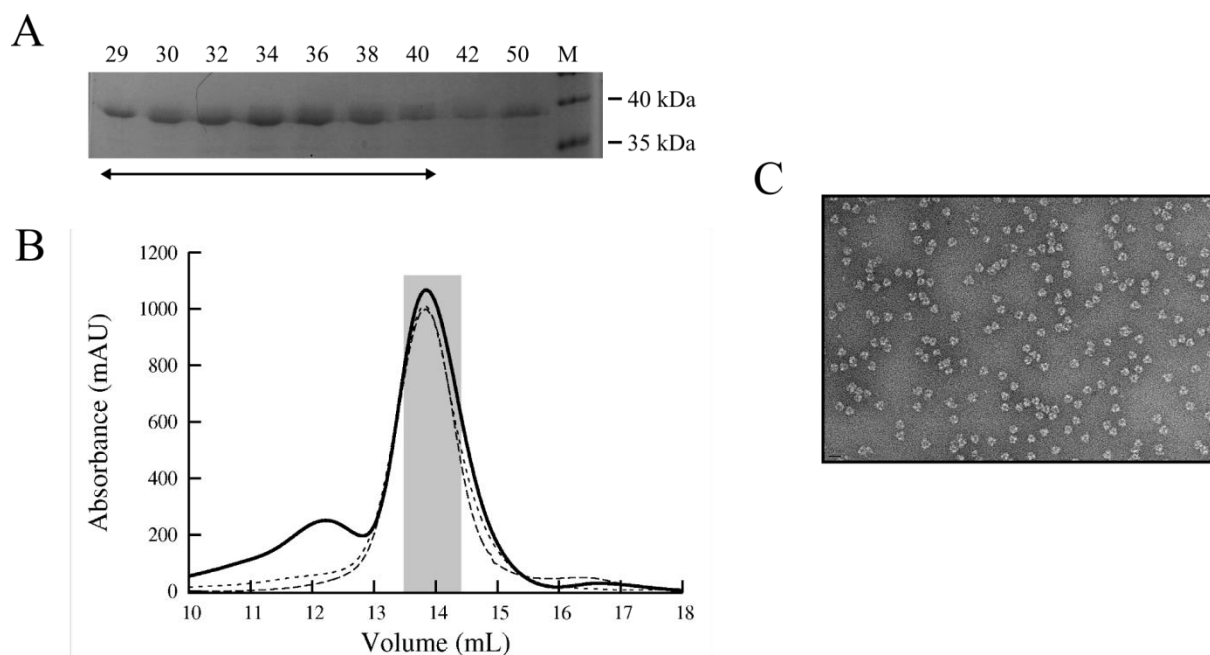
Ce protocole a permis d'obtenir 2.5 mg de protéine PhTET4 purifiée.

### 7.3 MICROSCOPIE ELECTRONIQUE

---

Afin de déterminer le type d'assemblage formé par le complexe purifié, un échantillon du pic obtenu à l'étape de gel filtration présenté en Figure 7-2B a été analysé en microscopie électronique. La préparation de l'échantillon est identique à celle de PhTET2 présenté en chapitre 5.

Les micrographies électroniques obtenues avec la protéine codée par la phase ouverte de lecture PH0737 montre qu'elle possède la structure tétraédrique creuse typique des aminopeptidases TET.



**Figure 7-2 : Purification et caractérisation de l'aminopeptidase PhTET4.** A) Fractions éluées en chromatographie échangeuses d'ions obtenues lors de la purification de la protéine PhTET4 analysées sur gel SDS-PAGE coloré au bleu de Coomassie. Les numéros des fractions sont indiqués en haut du gel. Les poids moléculaires correspondant aux bandes visibles du marqueur de poids moléculaire sur le gel sont indiqués à droite. Les fractions indiquées par une flèche ont été regroupées et analysées en colonne de gel filtration Superose 6 (voir B). M : Marqueur de poids moléculaire B) Profil d'élution des protéines purifiées en A analysées sur colonne de gel filtration Superose 6 (ligne épaisse). Les profils d'élution des dodécamères de PhTET2 et PhTET3 sont indiqués en lignes pointillés. La fraction du pic utilisée pour l'analyse en microscopie électronique est indiquée par une zone grisée. C) Micrographie électronique obtenues à partir de l'échantillon purifié en B. L'image montre une population homogène de complexe tétraédrique creux.

### 7.4 TEST ENZYMATIQUE

Après avoir confirmé que la conformation adoptée par la protéine purifiée est bien celle caractéristique des aminopeptidases TET, les propriétés enzymatiques de la protéine PhTET4 ont été étudiées.

Le protocole utilisé pour les études enzymatiques des aminopeptidases PhTET2 et PhTET3 étant le plus rapide à mettre en œuvre, c'est le premier qui ait été utilisé (voir section 2.1.2 et chapitre 4). Durant cette série de tests, j'ai tenté de détecter l'activité de PhTET4 sur 15 différents monoacyl-pNA, la liste des peptides utilisés est la suivante:

- Substrats spécifiques de PhTET1 : Asp-pNA, Ala-pNA
- Substrats spécifiques de PhTET2 : Leu-pNA, Met-pNA, Ile-pNA
- Substrats spécifiques de PhTET3 : Lys-pNA, Arg-pNA, Glu-pNA
- Substrats non hydrolysés par PhTET1, PhTET2 ou PhTET3 : His-pNA, Pro-pNA, Phe-pNA, Tyr-AMC et Trp-AMC

Aucune activité n'a été détectée en utilisant ces différents peptides substrats. Cependant, il a été observé durant la caractérisation de PhTET1 que l'enzyme était incapable d'hydrolyser les liaisons peptidiques de type X-pNA ou X-AMC (Schoenh 2006). L'hypothèse selon laquelle PhTET4 serait également incapable de prendre en charges ces peptides pourrait expliquer l'absence d'activité observée.

Une seconde stratégie a donc été mise en œuvre afin de tenter de détecter l'activité de PhTET4. L'enzyme a été incubée en présence d'hydrolysats de caséine, le mélange de peptides complexe est analysé par chromatographie en phase inverse sur un système HPLC afin de détecter une modification du profil chromatographique et donc une activité de la peptidase, une expérience identique a été décrite dans le chapitre 4 sur la figure 7. Cependant dans le cas de PhTET4, aucune activité hydrolytique n'a pu être observée. Ceci pourrait être expliqué par le protocole de préparation de l'hydrolysats de caséine. En effet, dans le cas où ce mélange serait issu de l'hydrolyse de la caséine par une seule enzyme, si l'activité de celle-ci est très spécifique d'un type d'un résidu, la variété des extrémités N-terminales serait alors faible. Il est alors possible qu'aucun résidu spécifique de l'activité de PhTET4 ne soit exposé pour son hydrolyse.

D'autres essais enzymatiques sont en cours dans le laboratoire afin d'identifier l'activité de PhTET4.

### 7.5 CRISTALLISATION

---

Un premier criblage pour les conditions de cristallisation a été réalisé à l'aide d'un robot de cristallisation nanogoutte à l'European Molecular Biology Laboratory (EMBL) de Grenoble. A la suite de ce criblage, les conditions ayant conduit à la formation de cristaux ont été reproduites et les gammes de concentration élargies sur des plaques à 24 puits en utilisant la technique de diffusion de vapeur. Une condition en particulier a conduit à la formation de cristaux qui ont été analysés par diffraction des rayons X.

Les cristaux analysés dont les jeux de données ont été retenus pour analyse ont été obtenus dans la condition suivante :  $\text{MgCl}_2$  0,1 M ; HEPES 0,1 M ; PEG400 22% ; pH= 8.

Plusieurs jeux de données ont été enregistrés, ceux-ci sont en cours d'analyse.

### 7.6 DISCUSSION ET PERSPECTIVES :

---

L'identification d'une quatrième TET chez *Pyrococcus horikoshii* sur la base d'une étude phylogénétique a conduit à la purification d'un complexe dodécamérique dont la structure observée en microscopie électronique montre une particule tétraédrique creuse typique des aminopeptidases TET. La caractérisation enzymatique et structurale de ce complexe est en cours dans le laboratoire. Les analyses de la séquence de cette protéine, sur la base des critères définis durant la thèse pour l'identification des complexes TET, montre que l'ensemble des structures très conservées entre PhTET1, PhTET2 et PhTET3 le sont moins chez PhTET4. Cette distance par rapport aux autres TET de *P. horikoshii* explique le temps supplémentaire qu'a demandé la découverte du quatrième membre de cette famille d'aminopeptidases. L'analyse de la structure de PhTET4 sera nécessaire pour continuer à affiner les critères déterminants pour l'identification d'une TET. Malgré le fait que ces séquences ne soient pas parfaitement conservées, les analyses en gel filtration suggèrent que la protéine est assemblée sous la forme d'un dodécamère, et les images de microscopie électronique montrent un édifice possédant la forme caractéristique du tétraèdre creux des TET. Cependant l'activité enzymatique de cette aminopeptidase putative n'a pas pu être encore déterminée. Des nouveaux tests sont en cours dans le laboratoire, l'utilisation d'autres substrats sans chromophore ou encore l'utilisation de peptides aux extrémités bloquées représente les pistes suivies actuellement. Les données de cristallographie aux rayons X semblent être de bonne qualité, le traitement de ces données est également en cours. La détermination de la structure tridimensionnelle de TET4 devrait permettre d'orienter la recherche de l'activité et des substrats de la peptidase. Ceci devrait aussi aider à mieux comprendre le rôle de cette famille de peptidases chez les Archées hyperthermophiles. Ces résultats, associés à ceux que donneront la poursuite de l'analyse phylogénétique des TET chez les Archées et les Bactéries, devrait permettre de mieux comprendre les relations entre les différentes TET de *Pyrococcus horikoshii*.

### 8 CONCLUSIONS GENERALES ET PERSPECTIVES

L'objectif principal de cette thèse était d'apporter des informations sur les relations entre oligomérisation et fonction chez les métallopeptidases de la famille M42.

L'analyse des structures cristallographiques de différentes peptidases de la famille M42 associée à l'étude du processus d'assemblage d'un mutant de la peptidase TET2 de *Pyrococcus horikoshii* a permis de montrer que la particule tétraédrique de 450kDa se forme à partir d'un dimère dont la structure est similaire à celle qu'il adopte au sein du dodécamère. Les études fonctionnelles montrent que l'assemblage permet d'augmenter l'activité de l'enzyme vis à vis de peptides longs. Ce mécanisme d'activation original n'a jamais été observé. Nos études *in vivo* indiquent que ce processus est régulé puisqu'on observe dans les cellules l'accumulation de dimères non assemblés. De plus, des expériences réalisées au laboratoire ELMA chez l'Archaea halophile *H. salinarum* montrent des changements importants de l'état d'oligomérisation de TET en fonction des conditions physiologiques. Ces études devraient être étendues aux différentes TET de *Pyrococcus horikoshii*. La recherche des mécanismes qui permettent de stabiliser le dimère *in cellulo* est un objectif de l'équipe. Ceci doit permettre de comprendre quels sont les signaux qui déclenchent l'assemblage des dodécamères *in vivo*. Une hypothèse est que les dimères représentent une forme de stockage de la protéine, rapidement mobilisable en fonction des besoins de la cellule, comme par exemple en conditions de carence nutritionnelle. Les études enzymatiques menées sur le dimère libre de TET montrent que celui-ci est dans une conformation active et est capable de cliver des peptides courts. Il pourrait donc être lui-même impliqué dans des processus physiologiques. La détermination du système de régulation de l'oligomérisation *in vivo* devrait permettre de mieux comprendre pourquoi cette forme active de la peptidase est accumulée dans la cellule.

Notre étude a également révélé le mécanisme d'assemblage du dodécamère de TET. Ce travail présente un intérêt général au delà de la thématique des peptidases TET. En effet on ne sait que peu de choses sur les déterminants structuraux et les processus biophysiques qui permettent d'assembler des édifices moléculaires complexes. Nous montrons ici que le processus est très contrôlé et procède via un intermédiaire



hexamérique. Plus encore, dans le cas de l'assemblage du complexe hétéro-oligomérique contenant des sous unités TET2 et TET3, nous montrons que le processus d'oligomérisation favorise la formation de poches catalytiques contenant les deux types de site actif. Ce mode de compartimentalisation "rationalisé" confère au complexe TET2-TET3 une efficacité de destruction des polypeptides supérieure à celle d'un mélange de complexes homo-oligomériques. Ainsi, on peut suggérer que les interfaces contiennent une information qui pilote la séquence d'assemblage en vue d'optimiser l'efficacité catalytique du complexe enzymatique. Là encore, l'étude phylogénomique des peptidases TET couplées à la caractérisation structurale et fonctionnelle de nouveaux complexes pourrait permettre d'aller plus loin dans la compréhension des mécanismes évolutifs conduisant à la formation des structures quaternaires et de la spécificité fonctionnelle des peptidases TET.

Nos études préliminaires montrent que l'histoire évolutive des TET chez les Archées a connu des événements multiples : acquisition de nouveaux systèmes à partir des Bactéries, perte ou conservation du système "ancestral", présence du système TET en multiples copies dans certains phylums, etc... Ceci pose de nombreuses questions sur le lien entre le rôle de TET dans les cellules et l'adaptation environnementale, notamment dans le cas des Archées marines ou retrouvées dans les sédiments profonds. La caractérisation de nouveaux systèmes TET, comme la nouvelle TET de *P.horikoshii* TET4 ou le système TET archéen, doit aussi permettre de savoir si la multiplicité des TET dans certaines cellules correspond à une optimisation du système, ou à l'acquisition de nouvelles fonctions à travers l'apparition de spécificités de substrats plus étroites. Ceci pourrait donner des indications sur le rôle physiologique des peptidases TET. Il serait également intéressant de savoir si les TET M42 procaryotes partagent un lien évolutif avec les TET M18 des eucaryotes.

Les informations obtenues au cours de ce travail nous permettent de progresser dans notre vision du rôle physiologique des peptidases TET. Ainsi il apparaît que dans le cas des TET de *Pyrococcus horikoshii*, TET2 et TET3 sont deux sous-unités d'un même complexe que nous avons nommé peptidasome. De par sa faible spécificité, ce complexe serait plutôt destiné à la dégradation complète des chaînes polypeptidiques qu'à des processus de maturation de peptides possédant une activité physiologique ou à une

fonction de modification des protéines néo-synthétisées. Le complexe TET1 apparaît quant à lui jouer un rôle physiologique différent. Il est destiné uniquement au clivage des extrémités N-ter comportant des acides glutamiques ou aspartiques. Il faut noter que les PepA et DNPEP, (les seuls complexes TET de bactérie et eucaryote pour lesquels une caractérisation fonctionnelle a été faite), sont aussi des aspartyl-aminopeptidases. Il semble donc que cette fonction soit conservée dans les 3 domaines du Vivant et jouera un rôle physiologique important. Quant au complexe TET4, nous ne connaissons pas ses substrats.

A ce stade des travaux, il devient important de réaliser des études génétiques: inactivation de gènes et surexpression. Cet outil est maintenant disponible tant chez les archées halophiles que chez certaines *Thermococcales* comme *P. furiosus* et *T. barophilus*. Un mutant knock-out viable d'une des deux TET de *H. volcanii* a été produit par C. Norais dans l'équipe de H. Myllykallio à l'Ecole polytechnique. L'étude des phénotypes de ce type de mutant, notamment par des analyses peptidomique et métabolomique, devrait nous renseigner sur le rôle et l'importance des peptidases TET dans le métabolisme.

### BIBLIOGRAPHIE

- Ando, S., K. Ishikawa, H. Ishida, Y. Kawarabayasi, H. Kikuchi & Y. Kosugi, (1999) Thermostable aminopeptidase from *Pyrococcus horikoshii*. *FEBS Letters* **447**: 25-28.
- Andrei, A.S., H.L. Banciu & A. Oren, (2012) Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. *FEMS Microbiol. Lett.* **330**: 1-9.
- Auld, D.S., (2004) Catalytic mechanisms for metallopeptidases. In: Handbook of Proteolytic Enzymes. A.J. Barrett, N.D. Rawlings & J.F. Woessner (eds). London: Academic Press,, pp. 268 289.
- Baek, D.H., J.J. Song, S.J. Kwon, C. Park, C.M. Jung & M.H. Sung, (2004) Characteristics of a new enantioselective thermostable dipeptidase from *Brevibacillus borstelensis* BCS-1 and its application to synthesis of a D-amino-acid-containing dipeptide. *Appl. Environ. Microbiol.* **70**: 1570-1575.
- Bar-Nun, S. & M.H. Glickman, (2012) Proteasomal AAA-ATPases: structure and function. *Biochim. Biophys. Acta.* **1823**: 67-82.
- Barrett, A.J. & J.K. McDonald, (1986) Nomenclature: protease, proteinase and peptidase. *Biochem. J.* **237**: 935.
- Baumeister, W., J. Walz, F. Zuhl & E. Seemuller, (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**: 367-380.
- Bennett, B. & R.C. Holz, (1997) Spectroscopically distinct cobalt(II) sites in heterodimetallic forms of the aminopeptidase from *Aeromonas proteolytica*: characterization of substrate binding. *Biochemistry* **36**: 9837-9846.
- Berger, A. & I. Schechter, (1970) Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **257**: 249-264.
- Borissenko, L. & M. Groll, (2005) Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J. Mol. Biol.* **346**: 1207-1219.
- Brandstetter, H., J.S. Kim, M. Groll & R. Huber, (2001) Crystal structure of the tricorn protease reveals a protein disassembly line. *Nature* **414**: 466-470.
- Breusing, N. & T. Grune, (2008) Regulation of proteasome-mediated protein degradation during oxidative stress and aging. *Biol. Chem.* **389**: 203-209.
- Brochier-Armanet, C., P. Forterre & S. Gribaldo, (2011) Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* **14**: 274-281.
- Brochier, C., S. Gribaldo, Y. Zivanovic, F. Confalonieri & P. Forterre, (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* **6**: R42.

- Bzymek, K.P. & R.C. Holz, (2004) The catalytic role of glutamate 151 in the leucine aminopeptidase from *Aeromonas proteolytica*. *J. Biol. Chem.* **279**: 31018-31025.
- Capes, M.D., P. DasSarma & S. DasSarma, (2012) The core and unique proteins of haloarchaea. *BMC Genomics* **13**: 39.
- Cavicchioli, R., (2011) Archaea--timeline of the third domain. *Nat. Mic. Reviews* **9**: 51-61.
- Cha, S.S., Y.J. An, C.R. Lee, H.S. Lee, Y.G. Kim, S.J. Kim, K.K. Kwon, G.M. De Donatis, J.H. Lee, M.R. Maurizi & S.G. Kang, Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *Embo. J.* **29**: 3520-3530.
- Chaban, B., S.Y. Ng & K.F. Jarrell, (2006) Archaeal habitats--from the extreme to the ordinary. *Can. J. Microbiol.* **52**: 73-116.
- Chapot-Chartier, M.P., F. Rul, M. Nardi & J.C. Gripon, (1994) Gene cloning and characterization of PepC, a cysteine aminopeptidase from *Streptococcus thermophilus*, with sequence similarity to the eucaryotic bleomycin hydrolase. *Eur. J. Biochem.* **224**: 497-506.
- Cheggour, A., L. Fanuel, C. Duez, B. Joris, F. Bouillenne, B. Devreese, G. Van Driessche, J. Van Beeumen, J.M. Frere & C. Goffin, (2000) The dppA gene of *Bacillus subtilis* encodes a new D-aminopeptidase. *Mol. Microbiol.* **38**: 504-513.
- Chen, G., T. Edwards, M. D'Souza V & R.C. Holz, (1997) Mechanistic studies on the aminopeptidase from *Aeromonas proteolytica*: a two-metal ion mechanism for peptide hydrolysis. *Biochemistry* **36**: 4278-4286.
- Chen, Y., L. Wu, R. Boden, A. Hillebrand, D. Kumaresan, H. Moussard, M. Baciú, Y. Lu & J. Colin Murrell, (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J.* **3**: 1093-1104.
- Cheng, T.C., V. Ramakrishnan & S.I. Chan, (1999) Purification and characterization of a cobalt-activated carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Protein. Sci.* **8**: 2474-2486.
- Chevrier, B., C. Schalk, H. D'Orchymont, J.M. Rondeau, D. Moras & C. Tarnus, (1994) Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure* **2**: 283-291.
- Chondrogianni, N. & E.S. Gonos, (2010) Proteasome function determines cellular homeostasis and the rate of aging. *Adv. Exp. Med. Biol.* **694**: 38-46.
- Chuang, C.K., B. Rockel, G. Seyit, P.J. Walian, A.-M. Schönege, J. Peters, P.H. Zwart, W. Baumeister & B.K. Jap, (2010) Hybrid molecular structure of the giant protease tripeptidyl peptidase II. *Nature Structural & Molecular Biology* **17**: 990-996.
- Clausen, T., C. Southan & M. Ehrmann, (2002) The HtrA family of proteases: implications for protein composition and cell fate. *Mol. Cell* **10**: 443-455.

- Colletier, J.P., A. Aleksandrov, N. Coquelle, S. Mraihi, E. Mendoza-Barbera, M. Field & D. Madern, (2012) Sampling the Conformational Energy Landscape of a Hyperthermophilic Protein by Engineering Key Substitutions. *Mol. Biol. Evol.* **29**: 1983-1994.
- Conway de Macario, E. & A.J. Macario, (2009) Methanogenic archaea in health and disease: a novel paradigm of microbial pathogenesis. *Int. J. Med. Microbiol.* **299**: 99-108.
- Cooper, E.M., C. Cutcliffe, T.Z. Kristiansen, A. Pandey, C.M. Pickart & R.E. Cohen, (2009) K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1. *Embo J.* **28**: 621-631.
- Coquelle, N., E. Fioravanti, M. Weik, F. Vellieux & D. Madern, (2007) Activity, stability and structural studies of lactate dehydrogenases adapted to extreme thermal environments. *J. Mol. Biol.* **374**: 547-562.
- Craik, C.S., M.J. Page & E.L. Madison, (2011) Proteases as therapeutics. *Biochem. J.* **435**: 1-16.
- Cunha, F.M., D.A. Berti, Z.S. Ferreira, C.F. Klitzke, R.P. Markus & E.S. Ferro, (2008) Intracellular peptides as natural regulators of cell signaling. *J. Biol. Chem.* **283**: 24448-24459.
- Daffonchio, D., S. Borin, T. Brusa, L. Brusetti, P.W. van der Wielen, H. Bolhuis, M.M. Yakimov, G. D'Auria, L. Giuliano, D. Marty, C. Tamburini, T.J. McGenity, J.E. Hallsworth, A.M. Sass, K.N. Timmis, A. Tselepidis, G.J. de Lange, A. Hubner, J. Thomson, S.P. Varnavas, F. Gasparoni, H.W. Gerber, E. Malinverno, C. Corselli, J. Garcin, B. McKew, P.N. Golyshin, N. Lampadariou, P. Polymenakou, D. Calore, S. Cenedese, F. Zanon & S. Hoog, (2006) Stratified prokaryote network in the oxic-anoxic transition of a deep-sea halocline. *Nature* **440**: 203-207.
- Dando, P.M., M. Fortunato, G.B. Strand, T.S. Smith & A.J. Barrett, (2003) Pyroglutamyl-peptidase I: cloning, sequencing, and characterisation of the recombinant human enzyme. *Protein Expr. Purif.* **28**: 111-119.
- De Paola, C.C., B. Bennett, R.C. Holz, D. Ringe & G.A. Petsko, (1999) 1-Butaneboronic acid binding to *Aeromonas proteolytica* aminopeptidase: a case of arrested development. *Biochemistry* **38**: 9048-9053.
- Delfosse, V., E. Girard, C. Birck, M. Delmarcelle, M. Delarue, O. Poch, P. Schultz & C. Mayer, (2009) Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family. *PloS one* **4**: e4712.
- DeLong, E.F. & N.R. Pace, (2001) Environmental diversity of bacteria and archaea. *Syst. Biol.* **50**: 470-478.
- Diaz-Perales, A., V. Quesada, J.R. Peinado, A.P. Ugalde, J. Alvarez, M.F. Suarez, F.X. Gomis-Ruth & C. Lopez-Otin, (2005) Identification and characterization of human archaemetzincin-1 and -2, two novel members of a family of metalloproteases widely distributed in Archaea. *J. Biol. Chem.* **280**: 30367-30375.

- Du, X., I.G. Choi, R. Kim, W. Wang, J. Jancarik, H. Yokota & S.H. Kim, (2000) Crystal structure of an intracellular protease from *Pyrococcus horikoshii* at 2-Å resolution. *PNAS* **97**: 14079-14084.
- Durá, M.A., V. Receveur-brechot, J.-p. Andrieu, C. Ebel, G. Schoehn, A. Roussel & B. Franzetti, (2005) Characterization of a TET-like Aminopeptidase Complex from the Hyperthermophilic Archaeon *Pyrococcus horikoshii* †. *Biochemistry* **44**: 3477-3486.
- Durá, M.A., E. Rosenbaum, A. Larabi, F. Gabel, F.M.D. Vellieux & B. Franzetti, (2009) The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol. Mic.* **72**: 26-40.
- Dutoit, R., N. Brandt, C. Legrain & C. Bauvois, (2012) Functional characterization of two M42 aminopeptidases erroneously annotated as cellulases. *PLoS One* **7**: e50639.
- Emsley, P., B. Lohkamp, W.G. Scott & K. Cowtan, (2010) Features and development of Coot. *Acta. Crystallogr. D Biol. Crystallogr.* **66**: 486-501.
- Falb, M., K. Muller, L. Konigsmaier, T. Oberwinkler, P. Horn, S. von Gronau, O. Gonzalez, F. Pfeiffer, E. Bornberg-Bauer & D. Oesterhelt, (2008) Metabolism of halophilic archaea. *Extremophiles* **12**: 177-196.
- Ferrer, M., O. Golyshina, A. Belouqui & P.N. Golyshin, (2007) Mining enzymes from extreme environments. *Curr. Opin. Microbiol.* **10**: 207-214.
- Forterre, P., (2013) The common ancestor of archaea and eukarya was not an archaeon. *Archaea* **2013**: 372396.
- Franzetti, B. & M.A. Durá, (2013) Chapter 371 – PhTET1 Aminopeptidase. In: Handbook of Proteolytic Enzymes. G.S. Neil D. Rawlings (ed). Oxford: Academic Press, pp. 1638–1645.
- Franzetti, B., G. Schoehn, D. Garcia, R.W.H. Ruigrok & G. Zaccai, (2002a) Characterization of the proteasome from the extremely halophilic archaeon. *Archaea* **1**: 53-61.
- Franzetti, B., G. Schoehn, J.-F. Hernandez, M. Jaquinod, R.W.H. Ruigrok & G. Zaccai, (2002b) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO* **21**: 2132-2138.
- Frottin, F., A. Martinez, P. Peynot, S. Mitra, R.C. Holz, C. Giglione & T. Meinnel, (2006) The proteomics of N-terminal methionine cleavage. *Mol. Cell Proteom.* **5**: 2336-2349.
- Geier, E., G. Pfeifer, M. Wilm, M. Lucchiari-Hartz, W. Baumeister, K. Eichmann & G. Niedermann, (1999) A giant protease with potential to substitute for some functions of the proteasome. *Science* **283**: 978-981.
- Godfroy, A., N.D. Raven & R.J. Sharp, (2000) Physiology and continuous culture of the hyperthermophilic deep-sea vent archaeon *Pyrococcus abyssi* ST549. *FEMS Microbiol. Lett.* **186**: 127-132.



- Goldberg, A.L., (2003) Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**: 895-899.
- Goldfarb, N.E. & B.M. Dunn, (2013) Chapter 45 - Human Immunodeficiency Virus 2 Retropepsin. In: Handbook of Proteolytic Enzymes. N.D.R. Salvesen (ed). Academic Press, pp. 199-204.
- Gonzales, T. & J. Robert-Baudouy, (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol. Rev.* **18**: 319-344.
- Gonzalez, J.M., Y. Masuchi, F.T. Robb, J.W. Ammerman, D.L. Maeder, M. Yanagibayashi, J. Tamaoka & C. Kato, (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2**: 123-130.
- Gonzalez, O., S. Gronau, M. Falb, F. Pfeiffer, E. Mendoza, R. Zimmer & D. Oesterhelt, (2008) Reconstruction, modeling & analysis of *Halobacterium salinarum* R-1 metabolism. *Mol. Biosyst.* **4**: 148-159.
- Graef, C., M. Schacherl, S. Waltersperger & U. Baumann, (2012) Crystal structures of archaeometzincin reveal a moldable substrate-binding site. *PLoS One* **7**: e43863.
- Gribaldo, S. & C. Brochier-Armanet, (2006) The origin and evolution of Archaea: a state of the art. *Philos Trans R Soc Lond B Biol Sci* **361**: 1007-1022.
- Gribaldo, S., P. Forterre & C. Brochier-Armanet, (2011) Archaea and the tree of life. *Res. Microbiol.* **162**: 1-4.
- Gribaldo, S., A.M. Poole, V. Daubin, P. Forterre & C. Brochier-Armanet, (2010) The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse? *Nat. Rev. Microbiol.* **8**: 743-752.
- Groll, M., M. Bochtler, H. Brandstetter, T. Clausen & R. Huber, (2005) Molecular machines for protein degradation. *ChemBiochem* **6**: 222-256.
- Guinier, A., (1939) "La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques.". *Ann. Phys.* **12**: 161-237.
- Haegeman, B., J. Hamelin, J. Moriarty, P. Neal, J. Dushoff & J.S. Weitz, (2013) Robust estimation of microbial diversity in theory and in practice. *ISME J.* **7**: 1092-1101.
- Halio, S.B., M.W. Bauer, S. Mukund, M. Adams & R.M. Kelly, (1997) Purification and Characterization of Two Functional Forms of Intracellular Protease PfpI from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* **63**: 289-295.
- Harroun, T.A., G.D. Wignall & J. Katsaras, (2006) Neutron Scattering for Biology. In: Neutron Scattering in Biology. Springer Berlin Heidelberg, pp. 1-18.
- Heinrich, J., K. Hein & T. Wiegert, (2009) Two proteolytic modules are involved in regulated intramembrane proteolysis of *Bacillus subtilis* RsiW. *Mol. Mic.* **74**: 1412-1426.

- Hofer, U., (2013) Environmental microbiology: Exploring diversity with single-cell genomics. *Nat. Rev. Microbiol.* **11**: 598.
- Holm, L. & P. Rosenstrom, (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**: W545-549.
- Humbard, M.A., H.V. Miranda, J.M. Lim, D.J. Krause, J.R. Pritz, G. Zhou, S. Chen, L. Wells & J.A. Maupin-Furlow, (2010) Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463**: 54-60.
- Ingmer, H. & L. Brondsted, (2009) Proteases in bacterial pathogenesis. *Res. Microbiol.* **160**: 704-710.
- Ishikawa, K., H. Ishida, Y. Koyama, Y. Kawarabayasi, J. Kawahara, E. Matsui & I. Matsui, (1998) Acylamino acid-releasing enzyme from the thermophilic archaeon *Pyrococcus horikoshii*. *J. Biol. Chem.* **273**: 17726-17731.
- Jacrot, B. & G. Zaccai, (1981) Determination of molecular weight by neutron scattering. *Biopolymers* **20**: 2413-2426.
- Jaenicke, R. & G. Bohm, (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**: 738-748.
- Jarrell, K.F., A.D. Walters, C. Bochiwal, J.M. Borgia, T. Dickinson & J.P. Chong, (2011) Major players on the microbial stage: why Archaea are important. *Microbiology* **157**: 919-936.
- Jiang, J., X. Zhang, Y. Chen, Y. Wu, Z.H. Zhou, Z. Chang & S.F. Sui, (2008) Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *PNAS* **105**: 11939-11944.
- Joshua-Tor, L., H.E. Xu, S.A. Johnston & D.C. Rees, (1995) Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. *Science* **269**: 945-950.
- Kambouris, N.G., D.J. Burke & C.E. Creutz, (1992) Cloning and characterization of a cysteine proteinase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**: 21570-21576.
- Kelly, R.M. & M.W. Adams, (1994) Metabolism in hyperthermophilic microorganisms. *Antonie Van Leeuwenhoek* **66**: 247-270.
- Kim, D., B.H. San, S.H. Moh, H. Park, D.Y. Kim, S. Lee & K.K. Kim, (2010) Structural basis for the substrate specificity of PepA from *Streptococcus pneumoniae*, a dodecameric tetrahedral protease. *Biochem. Biophys. Res. Commun.* **391**: 431-436.
- Kim, E., H. Kwak & K. Ahn, (2009) Cytosolic aminopeptidases influence MHC class I-mediated antigen presentation in an allele-dependent manner. *J. Immunol.* **183**: 7379-7387.
- Kim, J.S., M. Groll, H.J. Musiol, R. Behrendt, M. Kaiser, L. Moroder, R. Huber & H. Brandstetter, (2002) Navigation inside a protease: substrate selection and

- product exit in the tricorn protease from *Thermoplasma acidophilum*. *J. Mol. Biol.* **324**: 1041-1050.
- Krissinel, E. & K. Henrick, (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**: 774-797.
- Krojer, T., M. Garrido-Franco, R. Huber, M. Ehrmann & T. Clausen, (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* **416**: 455-459.
- Krojer, T., J. Sawa, E. Schafer, H.R. Saibil, M. Ehrmann & T. Clausen, (2008) Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**: 885-890.
- Krupovic, M., E.R. Quemin, D.H. Bamford, P. Forterre & D. Prangishvili, (2014) Unification of the globally distributed spindle-shaped viruses of the Archaea. *J. Virol.* **88**: 2354-2358.
- Lamm, O., (1929) Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* **21**: 1-4.
- Lefterov, I.M., R.P. Koldamova, M.I. Lefterova, D.R. Schwartz & J.S. Lazo, (2001) Cysteine 73 in bleomycin hydrolase is critical for amyloid precursor protein processing. *Biochem. Biophys. Res. Commun.* **283**: 994-999.
- Leigh, J.A., S.V. Albers, H. Atomi & T. Allers, (2011) Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol Rev.* **35**: 577-608.
- Lepage, P., M.C. Leclerc, M. Joossens, S. Mondot, H.M. Blottiere, J. Raes, D. Ehrlich & J. Dore, (2013) A metagenomic insight into our gut's microbiome. *Gut* **62**: 146-158.
- Lipinska, B., M. Zylicz & C. Georgopoulos, (1990) The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *J. Bact.* **172**: 1791-1797.
- Liu, Y., L.L. Beer & W.B. Whitman, (2012) Sulfur metabolism in archaea reveals novel processes. *Environ. Microbiol.* **14**: 2632-2644.
- Lloyd, K.G., L. Schreiber, D.G. Petersen, K.U. Kjeldsen, M.A. Lever, A.D. Steen, R. Stepanauskas, M. Richter, S. Kleindienst, S. Lenk, A. Schramm & B.B. Jorgensen, (2013) Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496**: 215-218.
- Marguet, E., M. Gaudin, E. Gaudiard, I. Fourquaux, S. le Blond du Plouy, I. Matsui & P. Forterre, (2013) Membrane vesicles, nanopods and/or nanotubes produced by hyperthermophilic archaea of the genus *Thermococcus*. *Biochem. Soc. Trans.* **41**: 436-442.
- Maupin-Furlow, J.A., (2013) Ubiquitin-like proteins and their roles in archaea. *Trends Microbiol.* **21**: 31-38.

- Maupin-Furlow, J.A., M.A. Humbard, P.A. Kirkland, W. Li, C.J. Reuter, A.J. Wright & G. Zhou, (2006) Proteasomes from structure to function: perspectives from archaea. *Curr. Top. Dev. Biol.* **75**: 125-169.
- Menyhard, D.K., A. Kiss-Szeman, E. Tichy-Racs, B. Hornung, K. Radi, Z. Szeltner, K. Domokos, I. Szamosi, G. Naray-Szabo, L. Polgar & V. Harmat, (2013) A self-compartmentalizing hexamer serine protease from *Pyrococcus horikoshii*: substrate selection achieved through multimerization. *J. Biol. Chem.* **288**: 17884-17894.
- Michaelis, L., M.L. Menten, K.A. Johnson & R.S. Goody, (2011) The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry* **50**: 8264-8269.
- Miller-Coleman, R.L., J.A. Dodsworth, C.A. Ross, E.L. Shock, A.J. Williams, H.E. Hartnett, A.I. McDonald, J.R. Havig & B.P. Hedlund, (2012) Korarchaeota diversity, biogeography, and abundance in Yellowstone and Great Basin hot springs and ecological niche modeling based on machine learning. *PLoS One* **7**: e35964.
- Montoya, S.E., C.E. Aston, S.T. DeKosky, M.I. Kamboh, J.S. Lazo & R.E. Ferrell, (1998) Bleomycin hydrolase is associated with risk of sporadic Alzheimer's disease. *Nat. Genet.* **18**: 211-212.
- Namba, Y., Y. Ouchi, A. Takeda, A. Ueki & K. Ikeda, (1999) Bleomycin hydrolase immunoreactivity in senile plaque in the brains of patients with Alzheimer's disease. *Brain Res.* **830**: 200-202.
- Namboori, S.C. & D.E. Graham, (2008) Acetamido sugar biosynthesis in the Euryarchaea. *J. Bacteriol.* **190**: 2987-2996.
- Nurun Nabi, A.H.M., F. Suzuki & T. Inagami, (2013) Chapter 11 - Renin. In: Handbook of Proteolytic Enzymes. N.D.R. Salvesen (ed). Academic Press, pp. 71-82.
- O'Farrell, P.a., F. Gonzalez, W. Zheng, S.a. Johnston & L. Joshua-Tor, (1999) Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease. *Structure* **7**: 619-627.
- O'Farrell, P.A. & L. Joshua-Tor, (2007) Mutagenesis and crystallographic studies of the catalytic residues of the papain family protease bleomycin hydrolase: new insights into active-site structure. *Biochem. J.* **401**: 421-428.
- Oger, P.M. & A. Cario, (2013) Adaptation of the membrane in Archaea. *Biophys. Chem.*
- Onoe, S., S. Ando, M. Ataka & K. Ishikawa, (2002) Active site of deblocking aminopeptidase from *Pyrococcus horikoshii*. *Biochem. Biophys. Res. Commun.* **290**: 994-997.
- Pikuta, E.V., R.B. Hoover & J. Tang, (2007) Microbial extremophiles at the limits of life. *Crit. Rev. Microbiol.* **33**: 183-209.
- Porciero, S., V. Receveur-Bréchet, K. Mori, B. Franzetti & A. Roussel, (2005) Expression, purification, crystallization and preliminary crystallographic analysis of a

- deblocking aminopeptidase from *Pyrococcus horikoshii*. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **61**: 239-242.
- Ratovitski, T., E. Chighladze, E. Waldron, R.R. Hirschhorn & C.A. Ross, (2011) Cysteine proteases bleomycin hydrolase and cathepsin Z mediate N-terminal proteolysis and toxicity of mutant huntingtin. *J. Biol. Chem.* **286**: 12578-12589.
- Rawlings, N.D., A.J. Barrett & A. Bateman, (2011) Asparagine peptide lyases: a seventh catalytic type of proteolytic enzymes. *J. Biol. Chem.* **286**: 38321-38328.
- Rawlings, N.D., A.J. Barrett & A. Bateman, (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **40**: D343-350.
- Rawlings, N.D., M. Waller, A.J. Barrett & A. Bateman, (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **42**: D503-509.
- Reeve, J.N. & C. Schleper, (2011) Archaea: very diverse, often different but never bad? *Curr. Opin. Microbiol.* **14**: 271-273.
- Remaut, H., C. Bompard-Gilles, C. Goffin, J.M. Frere & J. Van Beeumen, (2001) Structure of the *Bacillus subtilis* D-aminopeptidase DppA reveals a novel self-compartmentalizing protease. *Nat. Struct. Biol.* **8**: 674-678.
- Rockel, B., K.O. Kopec, A.N. Lupas & W. Baumeister, (2012) Structure and function of tripeptidyl peptidase II, a giant cytosolic protease. *Biochim. Biophys. Acta* **1824**: 237-245.
- Rohrwild, M., O. Cux, H.C. Huang, R.P. Moerschell, S.J. Yoo, J.H. Seol, C.H. Chung & A.L. Goldberg, (1996) HslV-HslU: A novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *PNAS* **93**: 5808-5813.
- Rosenbaum, E., M. Ferruit, M.A. Durá & B. Franzetti, (2011) Studies on the parameters controlling the stability of the TET peptidase superstructure from *Pyrococcus horikoshii* revealed a crucial role of pH and catalytic metals in the oligomerization process. *Biochimica et biophysica acta* **1814**: 1289-1294.
- Rosenbaum, E., F. Gabel, M.A. Dura, S. Finet, C. Clery-Barraud, P. Masson & B. Franzetti, (2012) Effects of hydrostatic pressure on the quaternary structure and enzymatic activity of a large peptidase complex from *Pyrococcus horikoshii*. *Arch. Biochem. Biophys.* **517**: 104-110.
- Russo, S. & U. Baumann, (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J. Biol. Chem.* **279**: 51275-51281.
- Samson, R.Y. & S.D. Bell, (2011) Cell cycles and cell division in the archaea. *Curr. Opin. Microbiol.* **14**: 350-356.
- Sato, T. & H. Atomi, (2011) Novel metabolic pathways in Archaea. *Curr. Opin. Microbiol.* **14**: 307-314.

- Sauer, R.T. & T.A. Baker, (2011) AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587-612.
- Schachman, H.K., (1957) Ultracentrifugation, Diffusion, and Viscometry. *Methods in Enzymology* **4**: 32-103.
- Schoehn, G., F.M.D. Vellieux, M. Asunción Durá, V. Receveur-Bréchet, C.M.S. Fabry, R.W.H. Ruigrok, C. Ebel, A. Roussel & B. Franzetti, (2006) An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron. *J. Biol. Chem.* **281**: 36327-36337.
- Schuck, P., (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**: 1606-1619.
- Schuhmann, H. & I. Adamska, (2012) Deg proteases and their role in protein quality control and processing in different subcellular compartments of the plant cell. *Physiologia Plantarum* **145**: 224-234.
- Schut, G.J., S.D. Brehm, S. Datta & M.W. Adams, (2003) Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. *J. Bacteriol.* **185**: 3935-3947.
- Sommaruga, S., E. Galbiati, J. Penaranda-Avila, C. Brambilla, P. Tortora, M. Colombo & D. Prosperi, (2014) Immobilization of carboxypeptidase from *Sulfolobus solfataricus* on magnetic nanoparticles improves enzyme stability and functionality in organic media. *BMC Biotechnol.* **14**: 82.
- Spiess, C., a. Beil & M. Ehrmann, (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**: 339-347.
- Stamper, C., B. Bennett, T. Edwards, R.C. Holz, D. Ringe & G. Petsko, (2001) Inhibition of the aminopeptidase from *Aeromonas proteolytica* by L-leucinephosphonic acid. Spectroscopic and crystallographic characterization of the transition state of peptide hydrolysis. *Biochemistry* **40**: 7035-7046.
- Stetter, K.O., (1999) Extremophiles and their adaptation to hot environments. *FEBS Lett.* **452**: 22-25.
- Sun, J., R.C. Hopkins, F.E. Jenney, P.M. McTernan & M.W. Adams, (2010) Heterologous expression and maturation of an NADP-dependent [NiFe]-hydrogenase: a key enzyme in biofuel production. *PLoS One* **5**: e10526.
- Sun, X., L. Peng, J. Guo, W. Chi, J. Ma, C. Lu & L. Zhang, (2007) Formation of DEG5 and DEG8 complexes and their involvement in the degradation of photodamaged photosystem II reaction center D1 protein in *Arabidopsis*. *The Plant cell* **19**: 1347-1361.
- Svergun, D.I., C. Barberato & M.H.J. Koch, (1995) CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.* **28**: 768-773.



- Svergun, D.I., M.H.J. Koch, P.A. Timmins & R.P. May, (2013) *Small angle x-ray and neutron scattering from solutions of biological macromolecules*, p. 358 pages.
- Svergun, D.I., S. Richard, M.H.J. Koch, Z. Sayers, S. Kuprin & G. Zaccai, (1998) Protein hydration in solution: experimental observation by X-ray and neutron scattering *PNAS* **95**: 768-773.
- Szeltner, Z., A.L. Kiss, K. Domokos, V. Harmat, G. Naray-Szabo & L. Polgar, (2009) Characterization of a novel acylaminoacyl peptidase with hexameric structure and endopeptidase activity. *Biochim. Biophys. Acta.* **1794**: 1204-1210.
- Takahashi, K., T. Taira, T. Niki, C. Seino, S.M. Iguchi-Ariga & H. Ariga, (2001) DJ-1 positively regulates the androgen receptor by impairing the binding of PIASx alpha to the receptor. *J. Biol. Chem.* **276**: 37556-37563.
- Takai, K. & K. Horikoshi, (1999) Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* **152**: 1285-1297.
- Talon, R., R. Kahn, M.A. Durá, O. Maury, F.M.D. Vellieux, B. Franzetti & E. Girard, (2011) Using lanthanoid complexes to phase large macromolecular assemblies. *J. Synchrotron Radiat.* **18**: 74-78.
- Tamura, N., F. Lottspeich, W. Baumeister & T. Tamura, (1998) The role of tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* **95**: 637-648.
- Tamura, T., N. Tamura, Z. Cejka, R. Hegerl, F. Lottspeich & W. Baumeister, (1996) Tricorn protease--the core of a modular proteolytic system. *Science* **274**: 1385-1389.
- Tehei, M., B. Franzetti, D. Madern, M. Ginzburg, B.Z. Ginzburg, M.T. Giudici-Orticoni, M. Bruschi & G. Zaccai, (2004) Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering. *EMBO Rep.* **5**: 66-70.
- Tehei, M., D. Madern, B. Franzetti & G. Zaccai, (2005) Neutron scattering reveals the dynamic basis of protein adaptation to extreme temperature. *J. Biol. Chem.* **280**: 40974-40979.
- Theobald, D.L. & P.A. Steindel, (2012) Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics* **28**: 1972-1979.
- Theriot, C.M., X. Du, S.R. Tove & A.M. Grunden, (2010a) Improving the catalytic activity of hyperthermophilic Pyrococcus prolidases for detoxification of organophosphorus nerve agents over a broad range of temperatures. *Appl. Microbiol. Biotechnol.* **87**: 1715-1726.
- Theriot, C.M., S.R. Tove & A.M. Grunden, (2010b) Characterization of two proline dipeptidases (prolidases) from the hyperthermophilic archaeon Pyrococcus horikoshii. *Appl. Microbiol. Biotechnol.* **86**: 177-188.

- Thiel, A., G. Michoud, Y. Moalic, D. Flament & M. Jebbar, (2014) Genetic manipulations of the hyperthermophilic piezophilic archaeon *Thermococcus barophilus*. *Appl. Microbiol. Biotechnol.* **80**: 2299-2306.
- Tomkinson, B., (1999) Tripeptidyl peptidases: enzymes that count. *Trends Biochem. Sci.* **24**: 355-359.
- Towne, C.F., I.A. York, L.B. Watkin, J.S. Lazo & K.L. Rock, (2007) Analysis of the role of bleomycin hydrolase in antigen presentation and the generation of CD8 T cell responses. *J. Immunol.* **178**: 6923-6930.
- Tsunasawa, S., (1998) Purification and application of a novel N-terminal deblocking aminopeptidase (DAP) from *Pyrococcus furiosus*. *J. Protein. Chem.* **17**: 521-522.
- Vairo Cavalli, S.E., M.C. Arribere, A. Cortadi, N.O. Caffini & N.S. Priolo, (2003) Morrenain b I, a papain-like endopeptidase from the latex of *Morrenia brachystephana* Griseb. (Asclepiadaceae). *J. Protein Chem.* **22**: 15-22.
- van Endert, P., (2011) Post-proteasomal and proteasome-independent generation of MHC class I ligands. *Cell Mol Life Sci* **68**: 1553-1567.
- Varshavsky, A., (2005) Regulated protein degradation. *Trends Biochem. Sci.* **30**: 283-286.
- Wagenfeld, A., J. Gromoll & T.G. Cooper, (1998) Molecular cloning and expression of rat contraception associated protein 1 (CAP1), a protein putatively involved in fertilization. *Biochem. Biophys. Res. Commun.* **251**: 545-549.
- Walz, J., a.J. Koster, T. Tamura & W. Baumeister, (1999) Capsids of tricorn protease studied by electron cryomicroscopy. *J. Struct. Biol.* **128**: 65-68.
- Wang, J., J.A. Hartling & J.M. Flanagan, (1997) The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**: 447-456.
- Ward, D.E., K.R. Shockley, L.S. Chang, R.D. Levy, J.K. Michel, S.B. Connors & R.M. Kelly, (2002) Proteolysis in hyperthermophilic microorganisms. *Archaea* **1**: 63-74.
- Welte, C. & U. Deppenmeier, (2011) Proton translocation in methanogens. *Methods Enzymol.* **494**: 257-280.
- Werner, F. & D. Grohmann, (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9**: 85-98.
- Whitman, W.B., F. Pfeifer, P. Blum & A. Klein, (1999) What archaea have to tell biologists. *Genetics* **152**: 1245-1248.
- Woese, C.R. & G.E. Fox, (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *PNAS* **74**: 5088-5090.
- Woese, C.R., O. Kandler & M.L. Wheelis, (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *PNAS* **87**: 4576-4579.
- Yao, T. & R.E. Cohen, (1999) Giant proteases: beyond the proteasome. *Curr. Biol.* **9**: R551-553.

- Ye, F. & M. Zhang, (2013) Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *Biochem. J.* **455**: 1-14.
- Zapras, A., T. Hoffmann, L. Stanek, K. Gunka, F.M. Commichau & E. Bremer, (2014) The gamma-aminobutyrate permease GabP serves as the third proline transporter of *Bacillus subtilis*. *J. Bacteriol.* **196**: 515-526.
- Zhan, D., A. Bai, L. Yu, W. Han & Y. Feng, (2014) Characterization of the PH1704 protease from *Pyrococcus horikoshii* OT3 and the critical functions of Tyr120. *PLoS One* **9**: e103902.
- Zheng, W., S.A. Johnston & L. Joshua-Tor, (1998) The unusual active site of Gal6/bleomycin hydrolase can act as a carboxypeptidase, aminopeptidase, and peptide ligase. *Cell* **93**: 103-109.
- Zierenberg, R.A., M.W. Adams & A.J. Arp, (2000) Life in extreme environments: hydrothermal vents. *PNAS* **97**: 12961-12962.
- Zimny, J., M. Sikora, A. Guranowski & H. Jakubowski, (2006) Protective mechanisms against homocysteine toxicity: the role of bleomycin hydrolase. *J. Biol. Chem.* **281**: 22485-22492.